For Reference

NOT TO BE TAKEN FROM THIS ROOM

For Reference

NOT TO BE TAKEN FROM THIS ROOM

Ex libris
universitatis
albertaensis

UNIVERSITY OF ALBERTA
QUAECUMQUE VERA

University of Alberta
Printing Department

Digitized by the Internet Archive
in 2019 with funding from
University of Alberta Libraries

https://archive.org/details/asymptoticdistri00jame

hesis
963 (F)
91

THE UNIVERSITY OF ALBERTA

ASYMPTOTIC DISTRIBUTIONS OF RESPONSE PROBABILITIES FOR A STOCHASTIC LEARNING MODEL

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE
OF MASTER OF SCIENCE

DEPARTMENT OF MATHEMATICS

by

JAMES VICTOR ZIDEK

EDMONTON, ALBERTA

AUGUST, 1963

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "Asymptotic Distributions of Response Probabilities for a Stochastic Learning Model", submitted by James Victor Zidek B.Sc. in partial fulfilment of the requirements for the degree of Master of Science.

ABSTRACT

This study is concerned with the limiting distributions which arise from a stochastic learning model proposed by Bush and Mosteller and discussed in "Stochastic Models for Learning" [1]. A summary of their work devoted to the class of learning processes with experimenter-controlled events is given in the first chapter of this thesis. For certain special cases of the Bush-Mosteller model, the limiting distributions have been found. The methods used in determining them are discussed in Chapter II. In the third chapter, some new results are obtained. For processes involving two experimenter-controlled events, a formula for the raw moments of the limiting distribution is obtained, and a method is developed by which the limiting distribution can be determined for any given member of a special class of these processes. This procedure is generalized to apply to processes with t-experimenter-controlled events $(t=2,3,\ldots)$.

ACKNOWLEDGEMENT

I am indebted to Dr. John Ross McGregor for suggesting to me the subject and for patiently assisting me during the preparation of this thesis.

TABLE OF CONTENTS

Page

CHAPTER I A STOCHASTIC LEARNING MODEL 1

Introduction 1

Section One - The Structure of the Model 2

Section Two - The Case of Two Experimenter-Controlled Events 9

Section Three - An Extension to t Experimenter-Controlled Events 20

CHAPTER II A REVIEW OF METHODS USED IN DETERMINING LIMITING DISTRIBUTIONS 25

Introduction 25

Section One - The Methods Used by Kemeny and Snell . . . 26

Section Two - The Methods Used by McGregor and Hui . . . 38

CHAPTER III ADDITIONAL RESULTS FOR CERTAIN LIMITING DISTRIBUTIONS . . 50

Introduction 50

Section One - Formulas For Moments 52

Section Two - The Case, $\alpha = 1/\sqrt[n]{2}$ $(n=1,2,\ldots)$ 61

Section Three - The Limit of the Sequence, $\{f_n(x)\}$ $(n=1,2,\ldots)$ 76

Section Four - The Case of t Experimenter-Controlled Events 78

BIBLIOGRAPHY 91

APPENDIX I 94

LIST OF TABLES

Page

TABLE 3.1 VALUES OF X_{3i} TOGETHER WITH THE DECIMAL APPROXIMATIONS, A_{3i}, $D_{31,i}$, AND $D_{32,i}$ OF X_{3i}, $T_{31}(X_{3i})$ AND $T_{32}(X_{3i})$, RESPECTIVELY $(i=1,2,3,4,5)$ 69

TABLE 3.2 VALUES OF X_{4i} TOGETHER WITH THE DECIMAL APPROXIMATIONS, A_{4i}, $D_{41,i}$ AND $D_{42,i}$ OF X_{4i}, $T_{41}(X_{4i})$ AND $T_{42}(X_{4i})$, RESPECTIVELY $(i=1,2,...9)$ 71

TABLES 3.3 VALUES OF X_{5i} TOGETHER WITH THE DECIMAL APPROXIMATIONS, A_{5i}, $D_{51,i}$ AND $D_{52,i}$ OF X_{5i}, $T_{51}(X_{5i})$ AND $T_{52}(X_{5i})$, RESPECTIVELY $(i=1,2,...,17)$ 73

LIST OF FIGURES

Page

FIGURE 1.1 VALUES OF $Q_i^n p$ PLOTTED AGAINST TRIAL NUMBER n 9

FIGURE 1.2 THE SUCCESSIVE SPLITS THAT A LARGE GROUP OF ANIMALS GOES THROUGH AFTER THREE EVENTS WHEN THE PROBABILITY OF APPLYING OPERATOR Q_1 IS π_1 , GIVEN AN OPERATOR IS TO BE APPLIED (EXPERIMENTER-CONTROLLED EVENTS). . . . 12

FIGURE 3.1 PROBABILITY DENSITY FUNCTION $f_n(x)$ FOR $n=1,2,3,4,5$ . . 75

CHAPTER I

A STOCHASTIC LEARNING MODEL

In a historical review of the mathematical theory of learning, Bush [2] suggests that the development of the present structure of the theory began in 1949 when several independent research programs were initiated. As a result of these investigations, several learning models were formulated and, of these, most are either of the "stimulus-sampling" or "stochastic-response" type. The first of these refers to models in which are incorporated a representation of the stimulus environment of an experimental subject and some means of describing the process by which environment influences behaviour. The assumptions made in such models lead to theorems concerning the way in which the response probability, as an index of learning, is transformed from one trial to another. Stochastic learning models assume a set of such transformation rules, and are concerned with predicting their consequences to learning, as measured by response probabilities. In Bush and Estes' "Studies in Mathematical Learning Theory" [3] both types are discussed and several specific models examined and compared.

One of the earliest learning models of the stochastic-response type is that proposed by Bush and Mosteller [4]. It is assumed that the transformation of response probabilities is carried out by the application of linear operators. Much of the work done jointly either by these authors alone [4], [5], [6] or with Thompson [7] is contained in their publication, "Stochastic Models for Learning".

The present chapter of this thesis provides an introduction to the learning model of Bush and Mosteller and summarizes those results which prove useful in later chapters. Section one describes the model and some

properties of the mathematical devices upon which it is based. A description of the specialized model which forms the subject of this thesis appears in section two. Finally, in section three, a generalization is suggested to which some of the results of the previous section are easily extended.

SECTION ONE

The model discussed in this thesis describes the learning process of an organism or subject in terms of the responses and outcomes occurring in each of a succession of trials in which the subject chooses to make exactly one of several responses to a given stimulus. The term, "learning", as it is used here, will refer to a systematic change in an organism's behavior.

In general, a set of r mutually exclusive and exhaustive response classes, $A_1, A_2, \ldots, A_r$, is considered and a trial defined as an opportunity for the selection of exactly one from among these r possible alternatives. For a given trial, there is associated with each A_i $(i = 1,2,\ldots,r)$ a probability, p_i, of its being chosen. Since these response classes are mutually exclusive and exhaustive,

$$\sum_{j=1}^{r} p_j = 1 \qquad \text{and} \qquad 0 \leq p_j \leq 1 \; . \tag{1.1.1}$$

It is assumed that on a given trial, a response, A_j, will have an outcome, O_k, from a set, $O_1, O_2, \ldots, O_s$, of mutually exclusive and exhaustive outcomes and that the occurrence of this response-outcome pair, (A_j,O_k), will alter, in a unique manner, the set of r response probabilities for the succeeding trial.

More generally, it may be assumed that the set of response probabilities, $p_1, p_2, \ldots p_r$, are changed from trial to trial in a unique way,

by the occurrence of one among a set of t events, $E_1, E_2, \ldots, E_t$, and that the way in which this change occurs is independent of earlier events in the succession of trials.

For our purposes, it will be sufficient to consider the case $r = s = 2$ in the subsequent discussion. The learning experiment, with this restriction, may be summarized as follows:

During a given trial, a subject selects one and only one of two responses A_1 and A_2 with probabilities p and $1 - p$, respectively. For either choice, exactly one of the outcomes O_1 and O_2 occurs. Finally, the probabilities of response, p and $1 - p$, are changed from trial to trial by the occurrence of events, $E_1, E_2, \ldots, E_t$ in a way which is independent of the previous occurrence of these events.

The model uses the response probabilities, p and $1 - p$, as an index of behavior and changes in behavior are represented by changes in these probabilities. We define the probability vector $\underline{p}$ as

$$\underline{p} = \begin{bmatrix} p \\ q \end{bmatrix}$$

where p and q are the probabilities of responses A_1 and A_2, respectively, and, by condition (1,1,1), $p + q = 1$. It is assumed that changes in these probabilities due to the occurrence of event $E_i (i = 1,2,\ldots,t)$ may be represented by the application of a two by two matrix operator, $\underline{T}_i$, to $\underline{p}$ where $\underline{T}_i$ is of the form

$$\underline{T}_i = \begin{bmatrix} u_{11,i} & u_{12,i} \\ u_{21,i} & u_{22,i} \end{bmatrix} \qquad (i=1,2,\ldots,t) \quad .$$

That is, the new vector of response probabilities is given by

$$\underline{T}_i\ \underline{p} = \begin{bmatrix} u_{11,i} & u_{12,i} \\ u_{21,i} & u_{22,i} \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix}$$

$$= \begin{bmatrix} u_{11,i}p + u_{12,i}q \\ u_{21,i}p + u_{22,i}q \end{bmatrix} \qquad (i=1,2,\ldots,t)\ .$$

The new probability of A_1 is $u_{11,i}p + u_{12,i}q$ and of A_2, $u_{21,i}p + u_{22,i}q$. By condition (1,1,1), therefore,

$$(u_{11,i}p + u_{12,i}q) + (u_{21,i}p + u_{22,i}q) = 1$$

or

$$(u_{11,i} + u_{21,i})p + (u_{12,i} + u_{22,i})q = 1\ .$$

Since this is true for all p and q satisfying $0 \leq p \leq 1$, $0 \leq q \leq 1$, and $p + q = 1$, then, in particular, for $p = 1$ and $q = 0$,

$$u_{11,i} + u_{21,i} = 1$$

and for $p = 0$, $q = 1$,

$$u_{12,i} + u_{22,i} = 1\ .$$

That is, $u_{11,i} = 1 - u_{21,i}$, $u_{22,i} = 1 - u_{12,i}$, and

$$\underline{T}_i = \begin{bmatrix} 1-u_{21,i} & u_{12,i} \\ u_{21,i} & 1-u_{12,i} \end{bmatrix} \qquad (i=1,2,\ldots,t)\ .$$

Let $b_i = u_{21,i}$ and $a_i = u_{12,i}$ so that

$$\underline{T}_i = \begin{bmatrix} 1-b_i & a_i \\ b_i & 1-a_i \end{bmatrix} \qquad (i=1,2,\ldots,t)\ .$$

Then, in this new notation,

$$\underline{T}_i\ \underline{p} = \begin{bmatrix} (1-b_i)p+a_iq \\ b_ip+(1-a_i)q \end{bmatrix} = \begin{bmatrix} Q_ip \\ \mathcal{Q}_iq \end{bmatrix} \quad ,$$

where

$$Q_i p = (1-b_i)p + a_i q$$

(1.1.2)

$$\mathcal{Q}_i q = b_i p + (1-a_i)q \qquad (i=1,2,\ldots,t) .$$

If we let $\alpha_i = 1 - a_i - b_i$ and $a_i = (1-\alpha_i)\lambda_i$ $(i=1,2,\ldots,t)$ then, by substitution and suitable rearrangements of the right hand side of equations (1.1.2), the following forms of $Q_i p$ are obtained:

a) SLOPE INTERCEPT FORM:

$$Q_i p = a_i + \alpha_i p \qquad (i=1,2,\ldots,t)$$

b) FIXED POINT FORM:

$$Q_i p = \alpha_i p + (1-\alpha_i)\lambda_i \qquad (i=1,2,\ldots,t) .$$

λ_i is the fixed point of this last transformation because $Q_i p = p$ if and only if $p = \lambda_i$. For either form,

(1.1.3) $$\mathcal{Q}_i q = 1 - Q_i p$$

by condition (1.1.1).

In order that $Q_i p$ $(i=1,2,\ldots,t)$ satisfy the conditions

$$0 \leq Q_i p \leq 1$$

for all p satisfying $0 \leq p \leq 1$, it is necessary to impose restrictions on the parameters a_i, b_i, α_i, and λ_i. Since $q = 1-p$ and, from equations (1.1.2), $Q_i p = a_i$ and $Q_i p = 1-b_i$ $(i=1,2,\ldots,t)$ when $p = 0$ and $p = 1$, respectively, the condition that $Q_i p$ be in the closed interval $[0,1]$ implies that

(1.1.4) $$0 \leq a_i \leq 1$$

and

(1.1.5) $$0 \leq b_i \leq 1 \qquad (i=1,2,\ldots,t) .$$

These conditions are sufficient since $Q_i p = (1-b_i)p + a_i q$ attains its maximum value when $b_i = 0$, $a_i = 1$, and $Q_i p = 1$, and its minimum value when $b_i = 1$, $a_i = 0$, and $Q_i p = 0$ $(i=1,2,\ldots,t)$.

Because $\alpha_i = 1-a_i-b_i$ by definition, it follows immediately from (1.1.4) and (1.1.5) that

$$-1 \leq \alpha_i \leq 1 \qquad (i=1,2,\ldots,t) .$$

If $\alpha_i > 0$, $Q_i p(=a_i+\alpha_i p)$ attains its maximum value when $p = 1$ and $Q_i p = a_i + \alpha_i$. This implies that

(1.1.6) $$\alpha_i \leq 1 - a_i \qquad (i=1,2,\ldots,t) .$$

On the other hand, if $\alpha_i < 0$, $Q_i p$ attains its minimum value when $p = 1$ and $Q_i p = a_i + \alpha_i$ and this implies that

(1.1.7) $$-a_i \leq \alpha_i \qquad (i=1,2,\ldots,t) .$$

Inequalities (1.1.6) and (1.1.7) combine to give

(1.1.8) $$-a_i \leq \alpha_i \leq 1 - a_i \qquad (i=1,2,\ldots,t) .$$

Finally, since λ_i is the fixed point of the transformation

$$Q_i p = \alpha_i p + (1-\alpha_i)\lambda_i \quad ,$$

and $0 \leq p \leq 1$, it is required that λ_i satisfy the conditions

(1.1.9) $$0 \leq \lambda_i \leq 1 \qquad (i=1,2,\ldots,t) .$$

In the subsequent discussion, the further restriction that

$$\alpha_i \geq 0 \qquad (i=1,2,\ldots,t) , \tag{1.1.10}$$

will be imposed to overcome practical difficulties associated with negative values of these parameters. For example, the oscillatory character of a sequence of response probabilities generated by successive applications of a given operator with a negative slope parameter is unrealistic in most learning experiments.

Condition (1.1.10) implies that $a_i + b_i \leq 1$ $(i=1,2,\ldots,t)$. The revised sets of restrictions are given below with their associated forms:

I SLOPE INTERCEPT FORM:

$$\begin{gathered} Q_i p = a_i + \alpha_i p \\ 0 \leq a_i \leq 1, \quad 0 \leq \alpha_i \leq 1-a_i \quad (i=1,2,\ldots,t) \end{gathered} \tag{1.1.11}$$

II FIXED POINT FORM:

$$\begin{gathered} Q_i p = \alpha_i p + (1-\alpha_i)\lambda_i \\ 0 \leq \lambda_i \leq 1, \quad 0 \leq \alpha_i \leq 1 \qquad (i=1,2,\ldots,t) . \end{gathered} \tag{1.1.12}$$

REPETITIVE APPLICATION OF A SINGLE OPERATOR, Q_i

In the fixed point form of equation (1.1.12), the operator (with the restrictions that $0 \leq \alpha_i \leq 1$ and $0 \leq \lambda_i \leq 1$) is applied to p whenever the event E_i occurs $(i=1,2,\ldots,t)$. After E_i occurs twice in succession, the probability of response A_1 is

$$\begin{aligned} Q_i^2 p = Q_i(Q_i p) &= \alpha_i[\alpha_i p + (1-\alpha_i)\lambda_i] + (1-\alpha_i)\lambda_i \\ &= \alpha_i^2 p + (1-\alpha_i^2)\lambda_i \quad (i=1,2,\ldots,t) . \end{aligned}$$

This suggests the induction hypothesis,

$$Q_i^n p = \alpha_i^n p + (1-\alpha_i^n)\lambda_i \qquad (i=1,2,\ldots,t) .$$

Under this hypothesis,

$$\begin{aligned} Q_i^{n+1} p &= Q_i(Q_i^n p) \\ &= \alpha_i(Q_i^n p) + (1-\alpha_i)\lambda_i \\ &= \alpha_i[\alpha_i^n p + (1-\alpha_i^n)\lambda_i] + (1-\alpha_i)\lambda_i \\ &= \alpha_i^{n+1} p + (1-\alpha_i^{n+1})\lambda_i . \end{aligned}$$

Therefore, if E_i occurs n times in succession, the probability that response A_1 occurs is

$$Q_i^n p = \alpha_i^n p + (1-\alpha_i^n)\lambda_i \qquad (i=1,2,\ldots,t) \tag{1.1.13}$$

after the last occurrence.

Now, when $0 \leq \alpha_i < 1$

$$\lim_{n \to \infty} Q_i^n p = \lambda_i \qquad (i=1,2,\ldots,t) ,$$

and, therefore, the fixed point λ_i is the limit point of the operator Q_i. In other words, if the same event (from among the set, $E_1, E_2, \ldots, E_t$) occurs repeatedly, the probability of response A_1 stabilizes at some fixed value. Moreover, equation (1.1.13) provides a means by which this response probability may be computed after a given number of repetitions of this event.

Figure 1.1 illustrates the effect of the repetitive application of an operator Q_i on p, where $p = 0.1$, $\lambda_i = 0.8$ and $\alpha_i = 0.9$.

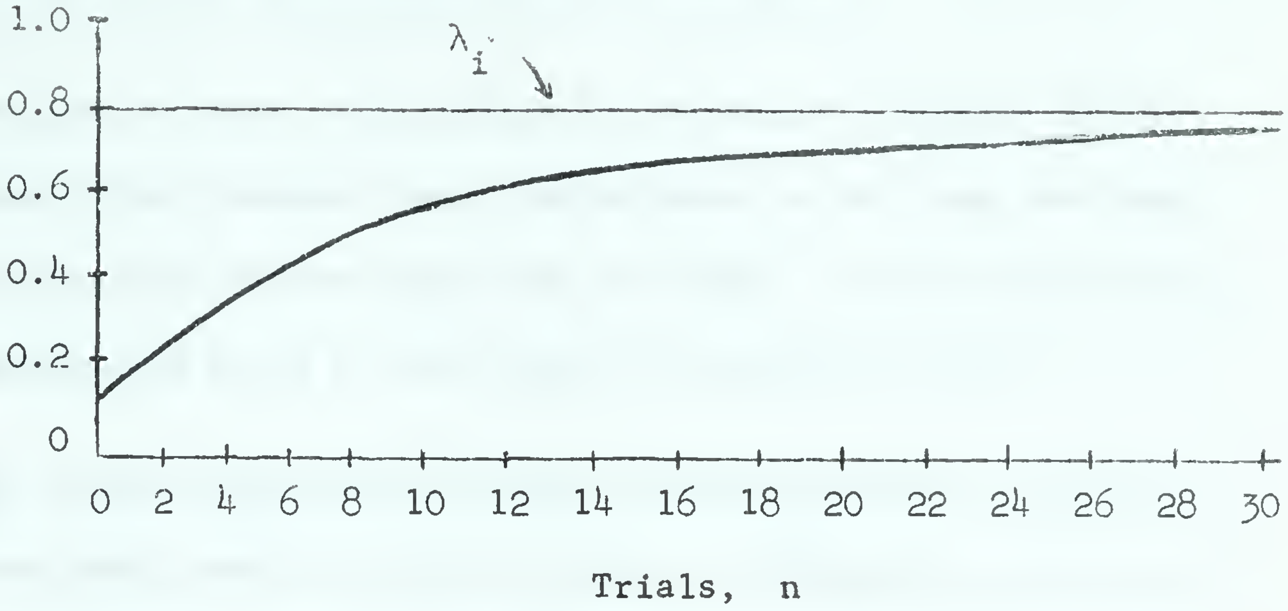

Fig. 1.1 Values of $Q_i^n p$ plotted against trial number n.
$Q_i^n p = \alpha_i^n p + (1-\alpha_i^n)\lambda_i$ with $p = 0.1$, $\lambda_i = 0.8$ and $\alpha_i = 0.9$ were used in plotting the curve.
(Reproduced from [1], page 60)

SECTION TWO

The model described above is applied to three categories of learning experiments, namely, those involving experimenter-controlled events, those with subject-controlled events, and those for which the events are experimenter-subject-controlled. The events $E_1, E_2, \ldots, E_t$ are said to be experimenter-controlled if they occur with fixed probabilities $\pi_1, \pi_2, \ldots, \pi_t$, respectively, chosen by the experimenter. If, in the case $t = 2$, we let the event E_1 be the choice of response A_1 and the event E_2 be the choice of response A_2, these events are subject-controlled. Finally, if outcomes O_1 and O_2 occur with fixed probabilities, π_1 and π_2, respectively, chosen by the experimenter, and we let E_{ij} be the occurrence of response A_i and outcome O_j $(i,j,=1,2)$, the events E_{11}, E_{12}, E_{21}, and E_{22} are called experimenter-subject-controlled events. Learning processes involving experimenter-controlled events are the subject of this thesis.

An example is provided by the Brunswick T-maze experiment.

Suppose that on each of a succession of trials a hungry rat is placed at the base of an elevated T-maze and allowed to run down the maze to a choice point where it either turns left or right. On each trial, food (the reward) is placed in a box at the right or left side of the T.

On each trial, then the rat is given the opportunity to choose between a right and left turn, and either choice is followed by exactly one of the two outcomes, reward and non-reward. Suppose it is assumed that:

(1) Turning right is rewarded a fixed proportion, π_1, of the trials when the rat went right and turning left is rewarded a fixed proportion, $\pi_2 = 1-\pi_1$, of the trials when the rat went left.

(2) A rewarded right turn and an unrewarded left turn have the same effect on the probability, p, of a right turn and so constitute an event class, E_1, with associated operator, Q_1.

(3) Similarly, a rewarded left turn and an unrewarded right turn constitute an event class, E_2, with associated operator, Q_2.

Then, the probability of E_1 is $p\pi_1 + (1-p)(1-\pi_2) = \pi_1$ and of E_2 is $(1-p)\pi_2 + p(1-\pi_1) = \pi_2$, so that the conditions of the general definition are met and this is an example of a learning experiment with two experimenter-controlled events.

PRE-ASYMPTOTIC DISTRIBUTION THEORY

Consider the case of two experimenter-controlled events, and hence, two operators, Q_1 and Q_2, which are applied with probabilities π_1 and π_2. Assume that $\pi_1 + \pi_2 = 1$ and let the organism's initial probability of choosing response A_1 be p_0. At the end of the first trial the probability

of response A_1 is Q_1p_0 with probability π_1 and Q_2p_0 with probability π_2, and after the second, is $Q_1^2p_0$ with probability π_1^2, $Q_2Q_1p_0$ with probability $\pi_2\pi_1$, $Q_1Q_2p_0$ with probability $\pi_1\pi_2$, and $Q_2^2p_0$ with probability π_2^2. This process carried through three trials is illustrated by a tree diagram in Figure 1.2.

If p_n denotes the probability of response A_1 on trial n, p_n has one of 2^n possible values, say, $p_{ni}(i=1,\ldots,2^n)$. Let

$$P_{ni} = P[p_n = p_{ni}] \qquad (i=1,2,\ldots,2^n).$$

Then, in terms of a large population of organisms, each having an initial probability p_0 of choosing response A_1, P_{ni} is the proportion of these organisms that would have a probability, p_{ni}, of choosing response A_1 after trial n $(i=1,2,\ldots,2^n;\ n=1,2,\ldots)$.

In the form of equation (1.1.11), $Q_1p_n = a_1+\alpha_1p_n$ and $Q_2p_n = a_2+\alpha_2p_n$, so that an organism with response probability p_{ni}, at trial n, has response probabilities $Q_1p_{ni} = a_1+\alpha_1p_{ni}$ and $Q_2p_{ni} = a_2+\alpha_2p_{ni}$ with conditional probabilities π_1 and π_2, respectively, at trial $n+1$ $(i=1,2,\ldots,2^n;\ n=1,2,\ldots)$. Thus, an organism at trial $n+1$ will have response probabilities Q_1p_{ni} and Q_2p_{ni} with probabilities π_1P_{ni} and π_2P_{ni}, respectively $(i=1,2,\ldots,2^n;n=1,2,\ldots)$.

Let $\bar{p}_n$ denote the expected probability of choosing response A_1 at trial n (In terms of a population of organisms, $\bar{p}_n$ is the average probability of response A_1 for this population on trial n). Then, with $\bar{p}_0 = p_0$,

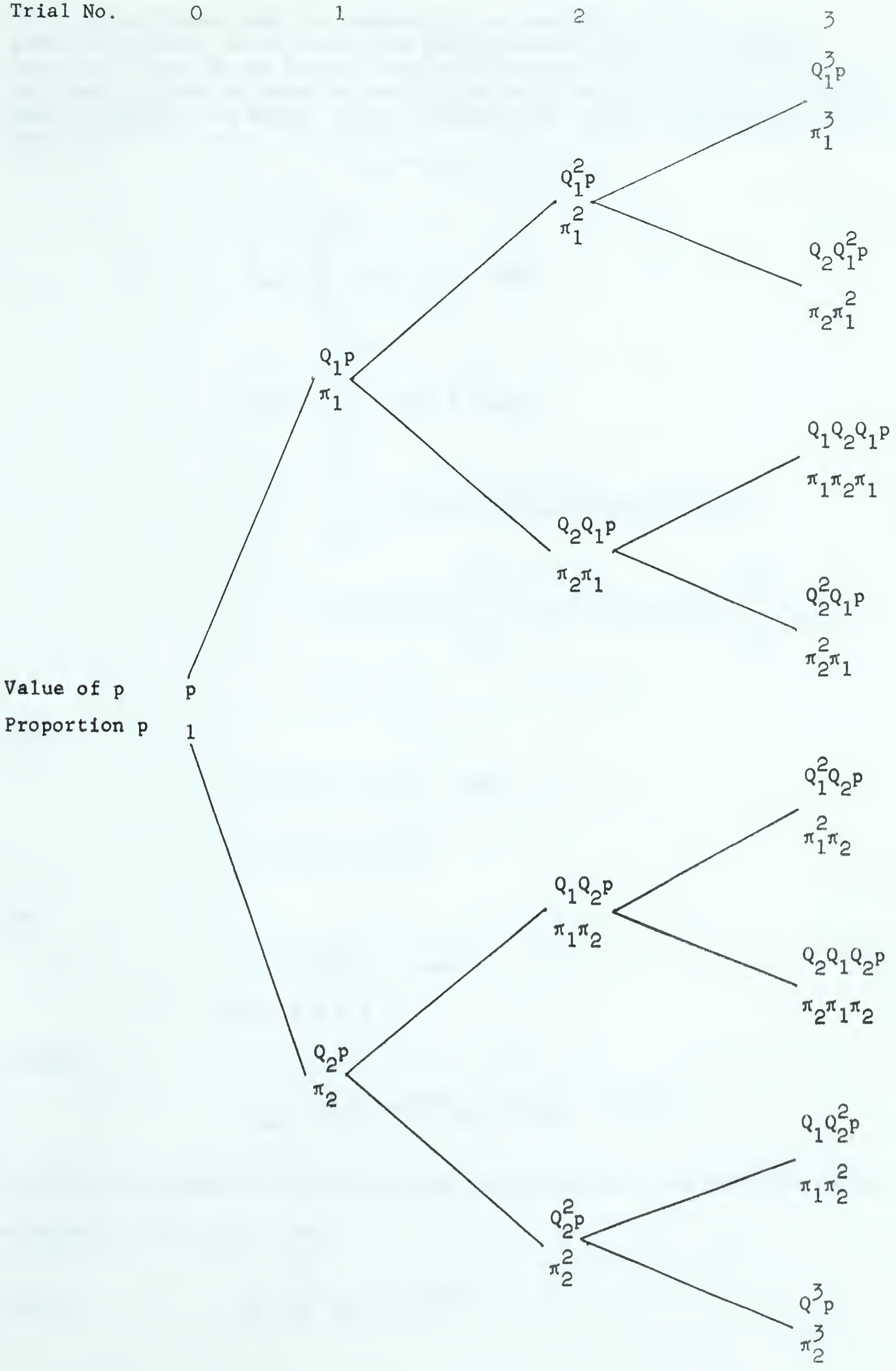

Fig. 1.2. The successive splits that a large group of animals goes through

after three events when the probability of applying operator Q_1 is π_1, given an operator is to be applied (experimenter-controlled events). The proportions in the groups, indicated beneath the p value of the group, have been written to parallel the Q's rather than the simplest form-thus, under $Q_2Q_1Q_2p$, we write $\pi_2\pi_1\pi_2$ rather than $\pi_1\pi_2^2$. (Reproduced from [1], page 70).

$$\bar{p}_n = \sum_{i=1}^{2^n} p_{ni}\, P_{ni} \;, \quad \text{and}$$

$$\begin{aligned}
\bar{p}_{n+1} &= \sum_{i=1}^{2^{n+1}} p_{n+1,i}\, P_{n+1,i} \\
&= \sum_{i=1}^{2^n} [(a_1+\alpha_1 p_{ni})\pi_1 P_{ni}+(a_2+\alpha_2 p_{ni})\pi_2 P_{ni}] \\
&= (\pi_1 a_1+\pi_2 a_2) \sum_{i=1}^{2^n} P_{ni} + (\pi_1\alpha_1+\pi_2\alpha_2) \sum_{i=1}^{2^n} p_{ni}P_{ni} \\
&= \bar{a} + \bar{\alpha}\, \bar{p}_n \;,
\end{aligned}$$

where

$$\bar{a} = \pi_1 a_1 + \pi_2 a_2 \quad \text{and}$$

$$\bar{\alpha} = \pi_1\alpha_1 + \pi_2\alpha_2 \;.$$

Let

$$\bar{a} = (1-\bar{\alpha})\bar{\lambda} \quad \text{and}$$

$$\bar{Q}p = \bar{a} + \bar{\alpha}\, p$$

so that

$$\bar{p}_{n+1} = \bar{Q}\, \bar{p}_n = \bar{Q}^{n+1} p_o = \bar{\alpha}\, \bar{p}_n + (1-\bar{\alpha})\bar{\lambda} \;.$$

A difference equation of this type has been solved and its solution given in equation (1.1.13). Thus,

$$\bar{p}_n = \bar{\alpha}^n p_o + (1-\bar{\alpha}^n)\bar{\lambda} \;. \tag{1.2.1}$$

MOMENTS

Let $V_{m,n}$ denote the m^{th} raw moment of the distribution of the response probability, p_n, where $P[p_n = p_{ni}] = P_{ni}$ and an organism on trial $n+1$ has response probabilities $Q_1 p_{ni}$ and $Q_2 p_{ni}$ with probabilities $\pi_1 P_{ni}$ and $\pi_2 P_{ni}$, respectively $(i=1,2,\ldots,2^n)$. Then,

$$V_{m,n} = \sum_{i=1}^{2^n} (p_{ni})^m P_{ni}, \; V_{0,n} = 1, \quad \text{and}$$

$$\begin{aligned} V_{m,n+1} &= \sum_{i=1}^{2^{n+1}} (p_{n+1,i})^m P_{n+1,i} \\ &= \sum_{i=1}^{2^n} [(a_1+\alpha_1 p_{ni})^m \pi_1 P_{ni} + (a_2+\alpha_2 p_{ni})^m \pi_2 P_{ni}] \\ &= \pi_1 \sum_{i=1}^{2^n} \sum_{k=0}^{m} \binom{m}{k} a_1^{m-k} \alpha_1^k (p_{ni})^k P_{ni} \\ &\quad + \pi_2 \sum_{i=1}^{2^n} \sum_{k=0}^{m} \binom{m}{k} a_2^{m-k} \alpha_2^k (p_{ni})^k P_{ni} \; . \end{aligned}$$

Thus,

$$V_{m,n+1} = \pi_1 \sum_{k=0}^{m} \binom{m}{k} a_1^{m-k} \alpha_1^k V_{k,n} + \pi_2 \sum_{k=0}^{m} \binom{m}{k} a_2^{m-k} \alpha_2^k V_{k,n} \; . \tag{1.2.2}$$

Thus, $V_{m,n+1}$ may be computed from the first $m+1$ raw moments, $V_{0,n},\ldots,V_{m,n}$, of the distribution of response probabilities on trial n.

In particular,

$$V_{1,n+1} = (\pi_1 a_1+\pi_2 a_2) + (\pi_1\alpha_1+\pi_2\alpha_2)V_{1,n} \quad \text{and} \tag{1.2.3}$$

(1.2.4) $$V_{2,n+1} = (\pi_1 a_1^2+\pi_2 a_2^2) + 2(\pi_1 a_1\alpha_1+\pi_2 a_2\alpha_2)V_{1,n} + (\pi_1\alpha_1^2+\pi_2\alpha_2^2)V_{2,n} = C_0 + C_1V_{1,n} + C_2V_{2,n} ,$$

where

(1.2.5) $$\begin{aligned} C_0 &= \pi_1 a_1^2 + \pi_2 a_2^2 \\ C_1 &= 2(\pi_1 a_1\alpha_1 + \pi_2 a_2\alpha_2) \\ C_2 &= \pi_1\alpha_1^2 + \pi_2\alpha_2^2 \end{aligned} .$$

Let

$$\bar{a} = \pi_1 a_1 + \pi_2 a_2$$

$$\bar{\alpha} = \pi_1\alpha_1 + \pi_2\alpha_2$$

$$\bar{Q}p = \bar{a} + \bar{\alpha}p ,$$

and define $V_{1,\infty}$ by the relation,

(1.2.6) $$V_{1,\infty} = \frac{\bar{a}}{1-\bar{\alpha}} ,$$

so that (1.2.3) may be expressed as

$$V_{1,n+1} = \bar{Q}V_{1,n} = \bar{Q}^{n+1}V_{1,0} = \bar{\alpha}V_{1,n} + (1-\bar{\alpha})V_{1,\infty} .$$

The solution of this difference equation is given in equation (1.1.13). Thus,

(1.2.7) $$V_{1,n} = \bar{\alpha}^n V_{1,0} + (1-\bar{\alpha}^n)V_{1,\infty} = V_{1,\infty} - (V_{1,\infty}- V_{1,0})\bar{\alpha}^n .$$

By substituting this result in (1.2.4), the following is obtained:

(1.2.8) $$V_{2,n+1} = (C_0+C_1V_{1,\infty}) - C_1(V_{1,\infty} - V_{1,0})\bar{a}^{n} + C_2V_{2,n} \,.$$

Let

$$C_0' = C_0 + C_1V_{1,\infty} \quad \text{and}$$

$$C_1' = -C_1(V_{1,\infty} - V_{1,0})$$

so that

(1.2.9) $$V_{2,n+1} = C_0' + C_1'\,\bar{\alpha}^{n} + C_2V_{2,n} \,.$$

Since

$$V_{2,0} = V_{1,0}^2 = p_0^2 \quad ,$$

$$V_{2,1} = C_0' + C_1' + C_2\,p_0^2 \quad .$$

Thus,

$$\begin{aligned} V_{2,2} &= C_0' + C_1'\,\bar{\alpha} + C_2(C_0' + C_1' + C_2\,p_0^2) \\ &= C_0'(1+C_2) + C_1'(\bar{\alpha}+C_2) + C_2^2\,p_0^2 \end{aligned}$$

and

$$\begin{aligned} V_{2,3} &= C_0' + C_1'\bar{\alpha}^2 + C_2[C_0'(1+C_2)+C_1'(\bar{\alpha}+C_2)+C_2^2p_0^2] \\ &= C_0'(1+C_2+C_2^2) + C_1'(\bar{\alpha}^2+C_2\bar{\alpha}+C_2^2) + C_2^3p_0^2 \quad . \end{aligned}$$

These results suggest the induction hypothesis,

(1.2.10) $$V_{2,n} = C_0'\sum_{u=0}^{n-1} C_2^u + C_1'\sum_{u=0}^{n-1}\bar{\alpha}^{n-1-u}\,C_2^u + C_2^np_0^2 \quad .$$

Assuming that (1.2.10) is true for $n = 1,2,\ldots,k$ we have, for $n = k+1$ [and using (1.2.9)],

$$V_{2,k+1} = C_0' + C_1' \bar{\alpha}^k + C_2[C_0' \sum_{u=0}^{k-1} C_2^u + C_1' \sum_{u=0}^{k-1} \bar{\alpha}^{k-1-u} C_2^u + C_2^k p_0^2]$$

$$= C_0' \sum_{u=0}^{k} C_2^u + C_1' \sum_{u=0}^{k} \bar{\alpha}^{k-u} C_2^u + C_2^{k+1} p_0^2 \quad ,$$

which is the result obtained when $n = k+1$ in (1.2.10). Hence, (1.2.10) is true by induction.

If $C_2 \neq 1$ and $\bar{\alpha} \neq C_2$,

$$\sum_{u=0}^{n-1} C_2^u = \frac{1-C_2^n}{1-C_2} \quad ,$$

$$\sum_{u=0}^{n-1} \bar{\alpha}^{n-1-u} C_2^u = \frac{\bar{\alpha}^n - C_2^n}{\bar{\alpha} - C_2} \quad ,$$

and hence,

$$V_{2,n} = C_0' \left(\frac{1-C_2^n}{1-C_2}\right) + C_1' \left(\frac{\bar{\alpha}^n - C_2^n}{\bar{\alpha} - C_2}\right) + C_2^n p_0^2$$

(1.2.11)

$$= (C_0 + C_1 V_{1,\infty}) \frac{1-C_2^n}{1-C_2} - C_1(V_{1,\infty} - V_{1,0}) \frac{\bar{\alpha}^n - C_2^n}{\bar{\alpha} - C_2} + C_2^n p_0^2 \ .$$

Since $\pi_1 + \pi_2 = 1$,

$$\bar{\alpha} = \pi_1 \alpha_1 + \pi_2 \alpha_2 \ ,$$

$$C_2 = \pi_1 \alpha_1^2 + \pi_2 \alpha_2^2 \quad ,$$

and, from (1.1.12),

$$0 \leq \alpha_i \leq 1 \quad (i=1,2) \ ,$$

it follows that

$$0 \leq C_2 < 1$$

$$0 \leq \bar{\alpha} \leq 1 \ ,$$

and, therefore, that $\lim_{n \to \infty} V_{2,n}$ exists. Let

$$V_{2,\infty} = \lim_{n \to \infty} V_{2,n}$$

and σ^2_∞ $(= V_{2,\alpha} - V^2_{1,\infty})$ be the limiting variance. Then,

$$V_{2,\infty} = \frac{C_0+C_1V_{1,\infty}}{1 - C_2} \quad \text{and}$$

$$\sigma^2_\infty = \frac{C_0+C_1V_{1,\infty}}{1 - C_2} - V^2_{1,\infty} \quad . \tag{1.2.12}$$

ASYMPTOTIC DISTRIBUTION THEORY

A sequence of distributions of response probabilities has been defined in association with the sequence of trials in a learning experiment with two experimenter-controlled events. The proof of the existence of a limiting distribution for this case and the "trapping theorem" stated below are due to Karlin [8]. Both results are quoted from [1], Page 99 (where "p values" refer to response probabilities) and the proofs are omitted.

> ASYMPTOTIC DISTRIBUTION THEOREM. For two response classes and experimenter-controlled events, it has been shown that an asymptotic distribution exists, that is, that the p value distributions on trials n and n+1 may be made as close to one another as desired by making n sufficiently large. Furthermore, the distribution is independent of the initial distribution of p values ....

Karlin [8], page 727 states that in a subsequent paper the theory of [8] will be adopted for more general models, where, for example, more response classes are allowed.

> TRAPPING THEOREM FOR $r=2$ .... consider two event operators having unique limit points and non-negative α's, that is, operators for which $0 \leq \alpha_i < 1$. The theorem asserts that if the p value of any sequence is ever between the two

> limit points the sequence forever remains there; this part of the theorem is intuitively obvious, for neither operator can take a p value outside the interval. The theorem also asserts that, when the sequence is outside the interval, it will later enter the interval with probability 1, provided that the probability of application of each operator is never zero. For more than two event operators with unique limit points and non-negative α's, the same statements apply to the interval between the largest and smallest limit points ....

References to this theorem will be made in the course of the subsequent discussion.

The first and second raw moments, as well as the variance, of the limiting distribution have been determined and are given in equations (1.2.6) and (1.2.12). These results and the problems to be discussed in the succeeding chapters are simplified when the equal alpha condition, which arises when a symmetry requirement is placed on the events E_1 and E_2 with operators $Q_1p = \alpha_1 p + (1-\alpha_1)\lambda_1$ and $Q_2p = \alpha_2 p + (1-\alpha_2)\lambda_2$, is imposed. If $\tilde{Q}_2$ is the complementary operator of Q_2 so that

$$\begin{aligned}\tilde{Q}_2 q &= 1 - Q_2 p\\ &= 1 - \alpha_2 p - (1-\alpha_2)\lambda_2\\ &= \alpha_2 q + (1-\alpha_2)(1-\lambda_2) ,\end{aligned}$$

we define this symmetry requirement in the following way: Q_1p is the same function of p as $\tilde{Q}_2 q$ is of q. This is true if and only if

$$\alpha_1 = \alpha_2$$

and

$$1 - \lambda_2 = \lambda_1$$

or, in other words, the symmetry requirement is met, if and only if, for operators Q_1 and Q_2, the slope parameters are equal and the limit points, complementary.

With the restriction that $\alpha_1 = \alpha_2 = \alpha$, (1.2.6), (1.2.5), and (1.2.12) become

$$V_{1,\infty} = \pi_1\lambda_1 + \pi_2\lambda_2 \,, \tag{1.2.13}$$

$$\left\{\begin{aligned} C_0 &= (1-\alpha)^2(\pi_1\lambda_1^2 + \pi_2\lambda_2^2) \\ C_1 &= 2\alpha(1-\alpha)V_{1,\infty} \\ C_2 &= \alpha^2 \,, \qquad \text{and} \end{aligned}\right. \tag{1.2.14}$$

$$\left\{\begin{aligned} V_{2,\infty} &= \frac{1}{1+\alpha}\,[(1-\alpha)(\pi_1\lambda_1^2+\pi_2\lambda_2^2) + 2\alpha V_{1,\infty}] \\ \sigma_\infty^2 &= \frac{1-\alpha}{1+\alpha}(\pi_1\lambda_1^2 + \pi_2\lambda_2^2 - V_{1,\infty}^2) \,, \end{aligned}\right. \tag{1.2.15}$$

respectively.

SECTION THREE

Consider a learning process with t experimenter-controlled events, $E_1, E_2, \ldots, E_t$ and associated operators, $Q_1, Q_2, \ldots, Q_t$ which are applied with fixed probabilities $\pi_1, \pi_2, \ldots, \pi_t$, respectively. Then, if p_n denotes the probability of response A_1 on trial n, p_n has one of t^n possible values, say, p_{ni} $(i=1,2,\ldots,t^n)$. Assume $\sum_{i=1}^{t} \pi_i = 1$ and let

$$P_{ni} = P[p_n = p_{ni}] \qquad (i=1,2,\ldots,t^n).$$

If $\bar{p}_n$ denotes the expected probability of choosing response A_1 at trial n, $\bar{p}_0 = p_0$,

$$\bar{p}_n = \sum_{i=1}^{t^n} p_{ni}\, P_{ni} \,, \qquad \text{and}$$

$$\bar{p}_{n+1} = \sum_{i=1}^{t^{n+1}} p_{n+1,i} \, P_{n+1,i}$$

$$= \sum_{j=1}^{t} \pi_j \sum_{i=1}^{t^n} (a_j+\alpha_j \, p_{ni}) \, P_{ni}$$

$$= \sum_{j=1}^{t} \pi_j a_j + \bar{p}_n \sum_{j=1}^{t} \pi_j \alpha_j$$

$$= \bar{a} + \bar{\alpha} \, \bar{p}_n \quad ,$$

where

$$\bar{a} = \sum_{j=1}^{t} \pi_j a_j \quad \text{and} \quad \bar{\alpha} = \sum_{j=1}^{t} \pi_j \alpha_j \quad .$$

The solution of this difference equation is given in (1.2.1) and hence,

$$\bar{p}_n = \bar{\alpha}^n \, p_0 + (1-\bar{\alpha}^n)\bar{\lambda} \quad , \tag{1.3.1}$$

where

$$\bar{\lambda} = \frac{\bar{a}}{1 - \bar{\alpha}} \quad .$$

If $V_{m,n}$ denotes the m^{th} raw moment of the distribution of the response probability p_n ,

$$V_{m,n} = \sum_{i=1}^{t^n} (p_{ni})^m \, P_{ni} \; , \quad V_{0,n} = 1, \quad \text{and}$$

$$V_{m,n+1} = \sum_{i=1}^{t^{n+1}} (p_{n+1,i})^m \, P_{n+1,i}$$

$$= \sum_{j=1}^{t} \pi_j \sum_{i=1}^{t^n} (Q_j p_{ni})^m \, P_{ni}$$

$$= \sum_{j=1}^{t} \pi_j \sum_{i=1}^{t^n} (a_j + \alpha_j p_{ni})^m \, P_{ni} \quad .$$

By using the binomial expansion and interchanging the order of summation over i and k, the following is obtained:

$$V_{m,n+1} = \sum_{j=1}^{t} \pi_j \sum_{k=0}^{m} \binom{m}{k} a_j^{m-k} \alpha_j^k V_{k,n} \quad . \tag{1.3.2}$$

In particular,

$$V_{1,n+1} = \sum_{j=1}^{t} \pi_j \, (a_j+\alpha_j V_{1,n})$$

$$= \bar{a} + \bar{\alpha} \, V_{1,n} \quad ,$$

where

$$\bar{a} = \sum_{j=1}^{t} \pi_j \, a_j$$

and

$$\bar{\alpha} = \sum_{j=1}^{t} \pi_j \, \alpha_j \quad .$$

This difference equation has been solved and its solution given by (1.2.7). Thus,

$$V_{1,n} = V_{1,\infty} - (V_{1,\infty} - V_{1,0})\bar{\alpha}^n \tag{1.3.3}$$

where

$$V_{1,\infty} = \frac{\bar{a}}{1 - \bar{\alpha}} \quad .$$

Also,

$$V_{2,n+1} = \sum_{j=1}^{t} \pi_j \, (a_j^2+2a_j\alpha_j V_{1,n}+\alpha_j^2 V_{2,n})$$

$$= \sum_{j=1}^{t} \pi_j \, a^2{}_j + 2\Big(\sum_{j=1}^{t} \pi_j a_j \alpha_j\Big)V_{1,n} + \Big(\sum_{j=1}^{t} \pi_j \alpha_j^2\Big)V_{2,n} \tag{1.3.4}$$

$$= C_0 + C_1 \, V_{1,n} + C_2 \, V_{2,n} \quad ,$$

where

$$C_0 = \sum_{j=1}^{t} \pi_j a_j^2 \quad ,$$

(1.3.5) $$C_1 = 2\sum_{j=1}^{t} \pi_j a_j \alpha_j \quad ,$$

$$C_2 = \sum_{j=1}^{t} \pi_j \alpha_j^2 \quad .$$

Substituting $V_{1,n} = V_{1,\infty} - (V_{1,\infty} - V_{1,0})\bar{\alpha}^n$ in (1.3.4) gives

$$V_{2,n+1} = (C_0 + C_1 V_{1,\infty}) - C_1(V_{1,\infty} - V_{1,0})\bar{\alpha}^n + C_2 V_{2,n}$$

with solution, given in (1.2.11),

(1.3.6) $$V_{2,n} = (C_0 + C_1 V_{1,\infty})\frac{1-C_2^n}{1-C_2} - C_1(V_{1,\infty} - V_{1,0})\frac{\bar{\alpha}^n - C_2^n}{\bar{\alpha} - C_2} + C_2^n p_0^2$$

provided that $C_2 \neq 1$ and $\bar{\alpha} \neq C_2$.

Since $\sum_{j=1}^{t} \pi_j = 1$,

$$\bar{\alpha} = \sum_{j=1}^{t} \pi_j \alpha_j \quad ,$$

$$C_2 = \sum_{j=1}^{t} \pi_j \alpha_j^2 \quad ,$$

and, from (1.1.12),

$$0 \leq \alpha_i \leq 1 \qquad (i=1,2,\ldots,t) \ ,$$

it follows that

$$0 \leq C_2 < 1$$

$$0 \leq \bar{\alpha} \leq 1$$

and hence, that $V_{2,\infty} = \lim_{n \to \infty} V_{2,n}$ exists and is given by

$$(1.3.7) \qquad V_{2,\infty} = \frac{C_0 + C_1 V_{1,\infty}}{1 - C_2} .$$

The limiting variance, $\sigma^2_\infty = (=V_{2,\infty} - V^2_{1,\infty})$ is

$$(1.3.8) \qquad \sigma^2_\infty = \frac{C_0 + C_1 V_{1,\infty}}{1 - C_2} - V^2_{1,\infty} .$$

CHAPTER II

A REVIEW OF METHODS USED IN DETERMINING LIMITING DISTRIBUTIONS

The explicit form of the limiting distributions of response probabilities for processes involving two experimenter-controlled events is known for only a few special cases in which the equal alpha condition has been imposed. Most of these known limiting distributions have been determined by Kemeny and Snell [9] and McGregor and Hui [10]. The methods devised by these authors and the results which have been obtained by them are reviewed in this chapter.

Section one is devoted to a study of the work of Kemeny and Snell who consider the case for which $\alpha_1 = \alpha_2 = \alpha$, $\lambda_1 = 1$ and $\lambda_2 = 0$; that is, the case involving two operators, $Q_1 p = \alpha p + (1-\alpha)$ and $Q_2 p = \alpha p$ [It will be shown in Section two that the limiting distribution may, by a suitable transformation, be made independent of λ_1 and λ_2]. For this case, it is shown that the response probability, p_n, is distributed like the quantity

$$S_n = \alpha^n p_o + \sum_{j=1}^{n} \epsilon_j (1-\alpha)\alpha^{j-1} ,$$

where p_o is the initial probability of response A_1 and the ϵ_j (j=1,2,...) are independently distributed random variables such that $\epsilon_j = 1$ or 0 with probability π_1 and π_2, respectively. Then, the limiting distribution is the same as the distribution of the quantity

$$S = \sum_{j=1}^{\infty} \epsilon_j \ (1-\alpha)\alpha^{j-1}$$

which is independent of p_o. By considering this random variable and a recursion relation for the raw moments of the limiting distribution, explicit forms for the limiting distribution are obtained in a number of cases. The characteristic function of the limiting distribution is shown to be

$$E(e^{itS}) = \prod_{j=1}^{\infty} [\pi_2 + \pi_1 e^{it(1-\alpha)\alpha^{j-1}}] \quad .$$

McGregor and Hui, whose work is summarized in section two, obtain the corresponding characteristic function for general λ_1 and λ_2 and show that by means of a transformation of the limiting variate, it may be made independent of λ_1 and λ_2. The characteristic function so obtained is inverted in special cases to yield explicit forms of the corresponding limiting distributions.

SECTION ONE

Consider the special case of a Bush-Mosteller learning process with two experimenter-controlled events, E_1 and E_2, for which $\lambda_1 = 1$, $\lambda_2 = 0$, and $\alpha_1 = \alpha_2 = \alpha \neq 1$. Let the probabilities of E_1 and E_2 be π_1 and π_2, respectively. Then, a probability, p, of response A_1, becomes $Q_1 p = \alpha p + (1-\alpha)$ and $Q_2 p = \alpha p$ with probabilities π_1 and π_2, respectively.

Let p_o be the initial probability of response A_1 and let the random variable p_n be the probability of response A_1 after trial n. Then p_1 is $\alpha p_o + (1-\alpha)$ or αp_o and p_2 is $\alpha^2 p_o + \alpha(1-\alpha) + (1-\alpha)$, $\alpha^2 p_o + (1-\alpha)$, $\alpha^2 p_o + \alpha(1-\alpha)$ or $\alpha^2 p_o$. It is clear that the possible values of p_n may be written as

$$\alpha^n p_o + \sum_{j=1}^{n} c_j(1-\alpha)\alpha^{j-1} , \tag{2.1.1}$$

where c_j is 0 or 1. Moreover, c_j is 1 if and only if Q_1 is applied on trial $n-j+1$ $(j=1,2,\ldots,n)$, so that the probability of a particular value of p_n, written in the form of (2.1.1), is $\pi_1^r \pi_2^{n-r}$ where r is the number of c_j which have value 1.

Thus, the distribution of p_n is the same as that of the random variable

$$S_n = \alpha^n p_o + \sum_{j=1}^{n} \epsilon_j(1-\alpha)\alpha^{j-1} \quad (n=1,2,\ldots) ,$$

where ϵ_j $(j=1,2,\ldots)$ are independent random variables such that ϵ_j is 1 or 0 with probabilities π_1 and π_2, respectively. The limiting distribution, whose existence is assured by the asymptotic distribution theorem, is the distribution of

$$S = \sum_{j=1}^{\infty} \epsilon_j(1-\alpha)\alpha^{j-1} . \tag{2.1.2}$$

Let F_n be the distribution of S_n. Then,

$$E[(S_{n+1})^j \mid S_n = x] = \pi_1[\alpha x + (1-\alpha)]^j + \pi_2[\alpha x]^j$$

and

$$\begin{aligned} E[(S_{n+1})^j] &= \pi_1 \int_0^1 [\alpha x+(1-\alpha)]^j \, dF_n + \pi_2 \int_0^1 [\alpha x]^j \, dF_n \\ &= \pi_1 \sum_{k=0}^{j} \binom{j}{k}\alpha^k(1-\alpha)^{j-k} E[S_n^k] + \pi_2\alpha^j E[S_n^j] . \end{aligned} \tag{2.1.3}$$

A recursion relation for the raw moments of the limiting distribution is obtained by letting n tend to infinity in (2.1.3). Thus,

$$E[S^j] = \pi_1 \sum_{k=0}^{j} \binom{j}{k} \alpha^k (1-\alpha)^{j-k} \, E[S^k] + \pi_2 \alpha^j \, E[S^j]$$

$$= \frac{\pi_1}{1-\alpha^j} \sum_{k=0}^{j-1} \binom{j}{k} \alpha^k (1-\alpha)^{j-k} \, E[S^k] \quad , \tag{2.1.4}$$

where $E[S^o] = 1$,

$$E[S^j] = \lim_{n \to \infty} E[S_n^j] \quad \text{for} \quad j = 0,1,2,\ldots \quad .$$

We observe that

$$\lim_{\alpha \to 1} E[S^j] = \pi_1 \sum_{k=0}^{j-1} \binom{j}{k} (j-k) \left[\lim_{\alpha \to 1} \frac{(1-\alpha)^{j-k-1}}{j\alpha^{j-1}} \right] \left[\lim_{\alpha \to 1} E[S^k] \right]$$

$$= \pi_1 \lim_{\alpha \to 1} E[S^{j-1}]$$

$$= \pi_1^j$$

and, hence, that the limiting distribution for $\alpha = 1$ is concentrated on the point, π_1 .

THE DISTRIBUTION OF S FOR THE CASE $\alpha < \frac{1}{2}$

It has been shown that S is independent of the initial response probability, $p_o (=S_o)$. For simplicity it will be assumed that $p_o = 0$ and, hence, that

$$S_n = \sum_{j=1}^{n} \epsilon_j (1-\alpha)\alpha^{j-1} \quad (n = 1,2,\ldots) \tag{2.1.5}$$

$$S_o = 0 \quad .$$

Suppose that the values of $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ are known and that the corresponding value of S_n is denoted by x. Assume $\alpha \neq 0$. Then, if

$$\epsilon_{n+1} = 1$$

$$S \geq x + (1-\alpha)\alpha^n$$

while, if $\epsilon_{n+1} = 0$, S cannot exceed the result obtained when $\epsilon_r = 1$ $(r = n+2, n+3, \ldots)$. That is,

$$S \leq x + \sum_{k=n+2}^{\infty} (1-\alpha)\alpha^{k-1}$$

$$= x + (1-\alpha) \left[\frac{1}{1-\alpha} - \frac{1-\alpha^{n+1}}{1-\alpha}\right]$$

$$= x + \alpha^{n+1} \quad .$$

Now, $x + \alpha^{n+1} < x + (1-\alpha)\alpha^n$ when α satisfies the conditions

$$0 < \alpha < \tfrac{1}{2} \quad .$$

Therefore, if $S_n = x$, S is outside an open interval,

$$(x + \alpha^{n+1}, x + (1-\alpha)\alpha^n) \quad , \tag{2.1.6}$$

of length $\alpha^n(1-2\alpha)$. It will now be shown that the intervals obtained in this way by considering all possible values for S_n, for all possible n, are disjoint and of total length 1.

Let S_{ni} $(i=1,2,\ldots,2^n)$, with $S_{ni} < S_{n,i+1}$ denote the 2^n possible values of S_n, each of which may be written as

$$\sum_{j=1}^{n} c_j(1-\alpha)\alpha^{j-1} \quad ,$$

where c_j is 0 or 1. Suppose S_{ni} and S_{nj}, with $S_{ni} > S_{nj}$, are given by

$$S_{ni} = \sum_{j=1}^{n} a_j(1-\alpha)\alpha^{j-1}$$

and

$$S_{nj} = \sum_{j=1}^{n} b_j(1-\alpha)\alpha^{j-1} .$$

Then, $a_j = b_j$ for $j = 1,2,\ldots,m'-1$ while $a_{m'} = 1$ and $b_{m'} = 0$ for some m' such that $1 \leq m' \leq n$. Consequently,

$$\begin{aligned} S_{ni} - S_{nj} &= \sum_{j=m'}^{n} a_j(1-\alpha)\alpha^{j-1} - \sum_{j=m'+1}^{n} b_j(1-\alpha)\alpha^{j-1} \\ &\geq \alpha^{m'-1}(1-\alpha) - \alpha^{m'}(1-\alpha^{n-m'}) \\ &= \alpha^{n} + \alpha^{m'-1}(1-2\alpha) . \end{aligned}$$

From this result, it is clear that

$$(2.1.7) \qquad |S_{ni} - S_{nj}| \geq \alpha^{n-1}(1-\alpha) \quad (i,j=1,2,\ldots,2^{n};\ i \neq j) ,$$

and, since $\alpha^{n-1}(1-\alpha) > \alpha^{n}(1-2\alpha)$, the intervals obtained for a given n are disjoint.

Assume that some pair of intervals are not disjoint. Then, there exist m, n, i and k, with $m > n$, $1 \leq i \leq 2^{n}$ and $1 \leq k \leq 2^{m}$, for which one of the following is true:

$$(1) \quad S_{ni} + \alpha^{n+1} \leq S_{mk} + \alpha^{m+1} \leq S_{ni} + \alpha^{n}(1-\alpha)$$

$$(2) \quad S_{ni} + \alpha^{n+1} \leq S_{mk} + \alpha^{m}(1-\alpha) \leq S_{ni} + \alpha^{n}(1-\alpha) .$$

If (1) is true,

$$(2.1.8) \qquad S_{ni} + \alpha^{n+1} - \alpha^{m+1} \le S_{mk} \le S_{ni} + \alpha^{n}(1-\alpha) - \alpha^{m+1} .$$

But S_{mk} can be written as

$$S_{mk} = S_{nt} + \sum_{j=n+1}^{m} c_j(1-\alpha)\alpha^{j-1} ,$$

where c_j is 0 or 1 and $1 \le t \le 2^n$. If $1 \le t < i$, we have, using (2.1.7),

$$(2.1.9) \qquad S_{mk} \le S_{ni} - \alpha^{n-1}(1-\alpha) + \alpha^{n} - \alpha^{m} .$$

Inequalities (2.1.8) and (2.1.9) imply that

$$\alpha^{n+1} - \alpha^{m+1} \le 2\alpha^{n} - \alpha^{n-1} - \alpha^{m}$$

or

$$(1-\alpha)\alpha^{n-1} + \alpha^{m} \le 0$$

which is false for $0 < \alpha \le 1$. If $t = i$, (2.1.8) becomes

$$(2.1.10) \qquad \alpha^{n+1} - \alpha^{m+1} \le \sum_{j=n+1}^{m} c_j(1-\alpha)\alpha^{j-1} \le \alpha^{n}(1-\alpha) - \alpha^{m+1} .$$

But, if $c_{n+1} = 1$,

$$\sum_{j=n+1}^{m} c_j(1-\alpha)\alpha^{j-1} \ge \alpha^{n}(1-\alpha) > \alpha^{n}(1-\alpha) - \alpha^{m+1} ,$$

while if $c_{n+1} = 0$,

$$\sum_{j=n+1}^{m} c_j(1-\alpha)\alpha^{j-1} \le \sum_{j=n+2}^{m} (1-\alpha)\alpha^{j-1} = \alpha^{n+1} - \alpha^{m} < \alpha^{n+1} - \alpha^{m+1}$$

for $0 < \alpha < 1$. Therefore $i < t \le 2^n$. But, in this case

$$S_{mk} \geq S_{ni} + \alpha^{n-1}(1-\alpha) > S_{ni} + \alpha^{n}(1-\alpha) - \alpha^{m+1}$$

which implies a contradiction to (1). Thus, there exist no m, n, i, and k for which (1) is true. A contradiction to assumption (2) is implied by similar arguments. Hence, the intervals of the form given in (2.1.6) obtained by considering all possible values of S_n, for all possible n, are disjoint.

For a given n, there are 2^n intervals of total length $(2\alpha)^n(1-2\alpha)$. Therefore, the total length of intervals obtained for $n = 0,1,2,\ldots$ is 1. Since the possible values of S which may be written as

$$\sum_{j=1}^{\infty} c_j(1-\alpha)\alpha^{j-1} \quad , \tag{2.1.11}$$

where c_j is 0 or 1, are outside these intervals, the limiting distribution is concentrated on a set of Lebesgue measure zero. In particular, when $\alpha = 1/3$, (2.1.11) becomes

$$\sum_{j=1}^{\infty} (2c_j/3^j) \tag{2.1.12}$$

and the limiting distribution is concentrated on the set of points whose ternary representation excludes the digit 1, that is, the Cantor set.

Suppose

$$x = \sum_{j=1}^{n} c_j(1-\alpha)\alpha^{j-1}$$

is a possible value for S_n. Then, if

$$S = x + \sum_{j=n+1}^{\infty} \epsilon_j(1-\alpha)\alpha^{j-1} \quad ,$$

$S \geq x$ and

$$S \leq x + \sum_{j=n+1}^{\infty} (1-\alpha)\alpha^{j-1} = x + \alpha^n .$$

That is, S is in the closed interval

$$[x, x + \alpha^n] . \tag{2.1.13}$$

Any other value of S_n differs from x by at least $\alpha^{n-1}(1-\alpha)$ which exceeds α^n, the length of the interval given by (2.1.13). Thus, the intervals obtained for all possible values of S_n are disjoint and S is in the interval given by (2.1.13) if and only if $\epsilon_j = c_j$ for $j = 1,2,\ldots n$. Therefore, if F is the distribution function of S,

$$F(x+\alpha^n) - F(x) = \pi_1^r \pi_2^{n-r} \qquad (n = 1,2,\ldots) . \tag{2.1.14}$$

where r is the number of c_j which have value 1. That the function, so determined is not absolutely continuous is shown below.

Recalling the definition of an absolutely continuous function we have:

> A function $f(x)$ is said to be absolutely continuous in an interval (a,b) if, given ϵ, we can find a δ such that $\sum_{\nu=1}^{n} |f(x_\nu+h_\nu) -f(x_\nu)| \leq \epsilon$ for every set of non-overlapping intervals $(x_\nu, x_\nu+h_\nu)$ such that $\sum_{\nu=1}^{n} h_\nu \leq \delta$
>
> (See, for example, Titchmarsh [11] page 364)

For arbitrary n, the intervals $[S_{ni}, S_{ni} + \alpha^n]$ $(i = 1,2,\ldots,2^n)$ form a set of non-overlapping intervals for which, using (2,1,14),

$$\sum_{i=1}^{2^n} |F(S_{ni}+\alpha^n) - F(S_{ni})| = \sum_{r=0}^{n} \binom{n}{r} \pi_1^r \pi_2^{n-r} \tag{2.1.15}$$

$$= 1$$

whereas

$$(2.1.16) \qquad \sum_{i=1}^{2^n} \alpha^n = (2\alpha)^n \quad .$$

Since this latter sum tends to 0 as n tends to infinity it can be made less than any prescribed δ by making n sufficiently large. However, there exists no δ for which

$$\sum_{i=1}^{2^n} |F(S_{ni} + \alpha^n) - F(S_{ni})| \leq \epsilon$$

when $\epsilon < 1$. Therefore F is not absolutely continuous.

THE CASE $\alpha = \frac{1}{2}$

When $\alpha = \frac{1}{2}$, (2.1.2) becomes

$$S = \sum_{j=1}^{\infty} (\epsilon_j/2^j) \quad .$$

The set of possible values for S are all numbers of the form

$$(2.1.17) \qquad x = \sum_{j=1}^{\infty} (c_j/2^j) \ ,$$

where c_j is 0 or 1. But every real number in the interval $[0,1]$ can be written in the form of (2.1.17) or, in other words, in the diadic form $c_1c_2c_3\ldots$.

Let $[a,b]$ be an interval with

$$a = c_1c_2\ldots c_n0$$

$$b = c_1c_2\ldots c_n1 \quad .$$

Since $\sum_{j=n+2}^{\infty} (a_j/2^j) \leq \frac{1}{2^{n+1}}$, this interval contains all values of S which are of diadic form

$$c_1 c_2 \ldots c_n \, 0 \, a_{n+2} a_{n+3} \ldots \; .$$

Therefore,

$$\begin{aligned} F(b) - F(a) &= P[a < S \leq b] \\ &= P[\epsilon_{n+1} = 0] \prod_{j=1}^{n} P[\epsilon_j = c_j] \\ &= \pi_1^r \; \pi_2^{n-r+1} \quad , \end{aligned}$$

where r is the number of c_j which have value 1. In particular, when $\pi_1 = \pi_2 = \frac{1}{2}$,

$$F(b) - F(a) = 1/2^{n+1}$$

while

$$b - a = 1/2^{n+1} \quad .$$

This determines F as the uniform distribution in this case. When $\pi_1 \neq \pi_2$, the resulting distribution is not absolutely continuous as shown in the following discussion.

The limiting distribution is the distribution of the random variable

$$S = \sum_{j=1}^{\infty} (\epsilon_j/2^j)$$

where ϵ_j is 1 or 0 with probability π_1 and π_2, respectively.

For any x in $[0,1]$, let $a_n(x)$ be the number of one's in the first n digits of the diadic representation of x. Given $\pi_1 = p$, let $M(p)$ be the set of S such that $\lim\limits_{n \to \infty} a_n(S)/n = p$, that is, $S \in M(p)$ if and only if this limit holds. Since $P[\epsilon_j=1] = p$, we have, by the law of large numbers, that $P[S \subset M(p)] = 1$. On the other hand, each distinct value of π_1 in $[0,1]$ determines a distinct set, $M(\pi_1)$ of values of S so that if $p_1, p_2, \ldots$ are different values of π_1,

$$L[\bigcup_m M(p_m)] = \sum_m L[M(p_m)] \quad ,$$

where L denotes the Lebesgue measure. In particular, since $\pi_1 = \frac{1}{2}$ determines the uniform distribution over $[0,1]$, $L[M(\frac{1}{2})] = 1$ and since $\bigcup_m M(p_m) \subset [0,1]$, it follows that $L[M(p_m)] = 0$, $p_m \neq \frac{1}{2}$. This implies that the limiting distribution function for $\pi_1 \neq \frac{1}{2}$ is constant almost everywhere and hence, it is not absolutely continuous.

THE DISTRIBUTION OF S FOR $\alpha > \frac{1}{2}$, $\pi_1 = \frac{1}{2}$

Let us define a sequence of α's as $\alpha_n = 1/\sqrt[n]{2}$, n=1,2,..., and consider the corresponding learning processes for $\pi_1 = \pi_2 = \frac{1}{2}$ [Wintner [12] has shown that for each member of the sequence just defined, the limiting distribution of response probabilities is absolutely continuous]. When $\alpha = 1/\sqrt[n]{2}$, (2.1.2) becomes

$$S = \sum_{j=1}^{\infty} \epsilon_j (1 - 1/\sqrt[n]{2})(1/\sqrt[n]{2})^{j-1}$$

$$= (1-1/\sqrt[n]{2}) \left[\sum_{k=0}^{\infty} \left(\epsilon_{nk+1}/2^k \right) + (1/\sqrt[n]{2}) \sum_{k=0}^{\infty} \left(\epsilon_{nk+2}/2^k \right) + \ldots \right.$$

$$\left. + (1/\sqrt[n]{2})^{n-1} \sum_{k=0}^{\infty} \left(\epsilon_{nk+n}/2^k \right) \right] .$$

After the ϵ's are suitably relabelled, this may be written as

$$S = \sum_{i=1}^{n} 2(1 - 1/\sqrt[n]{2})2^{-(i-1)/n}\left(\sum_{j=1}^{\infty} \epsilon_{ij}/2^j\right)$$

(2.1.18)

$$= \sum_{i=1}^{n} a_i X_i \quad ,$$

where ϵ_{ij} is 0 or 1 with probability $\frac{1}{2}$,

$$a_i = 2(1 - 1/\sqrt[n]{2})2^{-(i-1)/n} \quad ,$$

and X_i, $i=1,2, \ldots, n$ are independent random variables each of which has the uniform distribution on the interval $[0, 1]$.

The distribution, say, F_{α_n} of the random variable S given in result (2.1.18) is the convolution of the distributions of $a_i X_i$ $(i=1,2, \ldots,n)$. Thus, the characteristic function of F_{α_n} is easily obtained and is the product of the characteristic functions of the distributions of the summands. Kemeny and Snell state that F_{α_n} may be found by the inversion of its characteristic function and that it is piecewise a polynomial function.

In chapter III another method of determining F_{α_n} for given n is devised and several solutions exhibited.

THE CHARACTERISTIC FUNCTION OF S .

By result (2.1.2) the limiting variate may be expressed as

$$S = \sum_{j=1}^{\infty} \epsilon_j (1-\alpha)\alpha^{j-1}$$

where $\{\epsilon_j\}$ are independent random variables such that ϵ_j is 1 or 0 with probability π_1 and π_2, respectively. That is, S is the infinite

sum of independent random variables which are $(1-\alpha)\alpha^{j-1}$ and 0 with probabilities π_1 and π_2, respectively $(j=1,2, \ldots)$. Thus, the characteristic function of the distribution of S is the infinite product of the characteristic functions of the distributions of the summands. Hence,

$$E(e^{itS}) = \prod_{j=1}^{\infty} [\pi_2 + \pi_1 e^{it(1-\alpha)\alpha^{j-1}}] .$$

When $\pi_1 = \pi_2 = \frac{1}{2}$

$$E(e^{itS}) = \prod_{j=1}^{\infty} [\frac{1+e^{it(1-\alpha)\alpha^{j-1}}}{2}]$$

$$= \prod_{j=1}^{\infty} e^{it(1-\alpha)\alpha^{j-1}/2} \cos[t(1-\alpha)\alpha^{j-1}/2]$$

$$= \exp(it/2) \prod_{j=1}^{\infty} \cos[t(1-\alpha)\alpha^{j-1}/2] .$$

SECTION TWO

As before, let p_n denote the probability of response A_1 on trial n, p_{ni} $(1,2, \ldots, 2^n)$ denote the 2^n possible values of p_n, and $P_{ni} = P[p_n = p_{ni}]$ $(1,2, \ldots, 2^n)$. Let $m_n(t) = E\{\exp(tp_n)\}$ be the moment generating function of the distribution of p_n. Then,

$$m_n(t) = \sum_{i=1}^{2^n} e^{tp_{ni}} P_{ni} .$$

Since p_{n+1} is $Q_1p_{ni} = a_1 + \alpha_1 p_{ni}$ or $Q_2p_{ni} = a_2 + \alpha_2 p_{ni}$ with probability $\pi_1 P_{ni}$ and $\pi_2 P_{ni}$, respectively,

$$m_{n+1}(t) = \sum_{i=1}^{2^n} \left[e^{t(a_1+\alpha_1 p_{ni})} \pi_1 P_{ni} + e^{t(a_2+\alpha_2 p_{ni})} \pi_2 P_{ni}\right]$$

$$= \pi_1 e^{ta_1} \sum_{i=1}^{2^n} e^{t\alpha_1 p_{ni}} p_{ni} + \pi_2 e^{ta_2} \sum_{i=1}^{2^n} e^{t\alpha_2 p_{ni}} p_{ni}$$

$$= \pi_1 e^{ta_1} m_n(t\,\alpha_1) + \pi_2 e^{ta_2} m_n(t\,\alpha_2) \; .$$

By letting $n \to \infty$, a functional equation for the moment generating function of the asymptotic distribution is obtained. It is

$$m(t) = \pi_1 e^{ta_1} m(t\,\alpha_1) + \pi_2 e^{ta_2} m(t\,\alpha_2) \; , \tag{2.2.1}$$

where $m(t) = \lim_{n \to \infty} m_n(t)$. When $\alpha_1 = \alpha_2 = \alpha$, (2.2.1) becomes

$$m(t) = m(\alpha t)\,[\pi_1 e^{ta_1} + \pi_2 e^{ta_2}] \; . \tag{2.2.2}$$

This equation implies that

$$m(\alpha t) = m(\alpha^2 t)\,[\pi_1 e^{t\alpha a_1} + \pi_2 e^{t\alpha a_2}]$$

giving

$$m(t) = m(\alpha^2 t)[\pi_1 e^{ta_1} + \pi_2 e^{ta_2}][\pi_1 e^{t\alpha a_1} + \pi_2 e^{t\alpha a_2}] \; .$$

Provided that the resulting infinite product converges, this procedure may be applied iteratively to yield the explicit solution

$$m(t) = \prod_{n=0}^{\infty} [\pi_1 e^{ta_1\alpha^n} + \pi_2 e^{ta_2\alpha^n}] \; . \tag{2.2.3}$$

That this product does converge is shown in the following.

Assume that $0 \le \alpha \le 1$, $a_1 \ge a_2$ and t is real. Then, using natural logarithims,

$$|\log m(t)| = \Big|\sum_{n=0}^{\infty} \log\,(\pi_1 e^{ta_1\alpha^n} + \pi_2 e^{ta_2\alpha^n})\Big|$$

When $-\infty < t < 0$, $t = -|t|$, and

$$
\begin{aligned}
|\log m(t)| &\le \sum_{n=0}^{\infty} |\log (\pi_1 e^{-|t|a_1\alpha^n} + \pi_2 e^{-|t|a_2\alpha^n})| \\
&\le \sum_{n=0}^{\infty} |\log (\pi_1 e^{|t|a_2\alpha^n} + \pi_2 e^{|t|a_1\alpha^n})| \\
&\quad + \sum_{n=0}^{\infty} \log [e^{|t|(a_1+a_2)\alpha^n}] \\
&\le \sum_{n=0}^{\infty} |t|a_1\alpha^n + \sum_{n=0}^{\infty} |t|(a_1+a_2)\alpha^n \\
&= \frac{|t|a_1}{1-\alpha} + \frac{|t|(a_1+a_2)}{1-\alpha} .
\end{aligned}
$$

When $0 < t < \infty$, $t = |t|$, and

$$
\begin{aligned}
|\log m(t)| &\le \sum_{n=0}^{\infty} |\log (\pi_1 e^{|t|a_1\alpha^n} + \pi_2 e^{|t|a_2\alpha^n})| \\
&\le \sum_{n=0}^{\infty} |t|a_1\alpha^n \\
&= \frac{|t|a_1}{1-\alpha} .
\end{aligned}
$$

In either case, $|\log m(t)|$ is finite and, hence, $m(t)$ converges for $-\infty < t < \infty$.

If the random variable of the limiting distribution is denoted by p , we know from the trapping theorem [see Bush and Mosteller [1], page 99] that $P[p \subset I] = 0$ if I is any interval which is disjoint

from the closed interval joining $\lambda_1(=\frac{a_1}{1-\alpha})$ and $\lambda_2 = (\frac{a_2}{1-\alpha})$. That is, the limiting response probabilities are distributed between the limit points λ_1 and λ_2. Let

$$x = [\frac{\lambda_2 + \lambda_1}{\lambda_2 - \lambda_1}] - [\frac{2p}{\lambda_2 - \lambda_1}]$$

so that x is distributed on the interval $(-1, 1)$. Let $m_x(t)$ denote the moment generating function of the distribution of x. Then,

$$\begin{aligned} m_x(t) &= E[\exp(tx)] \\ &= E\left[\exp([\frac{\lambda_2 + \lambda_1}{\lambda_2 - \lambda_1}]t - [\frac{2pt}{\lambda_2 - \lambda_1}])\right] \end{aligned}$$

or

$$m_x(t) = \exp([\frac{\lambda_2 + \lambda_1}{\lambda_2 - \lambda_1}]t)\ m\ (\frac{-2t}{\lambda_2 - \lambda_1}) \tag{2.2.4}$$

Thus,

$$m(t) = \exp[(\lambda_2 + \lambda_1)\ \frac{t}{2}]\ m_x\ [-\frac{t}{2}(\lambda_2 - \lambda_1)] .$$

This result may be substituted into equation (2.2.3) to obtain the explicit solution for $m_x(t)$. Thus,

$$m_x[-\frac{t}{2}(\lambda_2-\lambda_1)] = \exp[-\frac{t}{2}(\lambda_2+\lambda_1)] \prod_{n=0}^{\infty} [\pi_1 e^{t(1-\alpha)\alpha^n\lambda_1} + \pi_2 e^{t(1-\alpha)\alpha^n\lambda 2}]$$

or

$$m_x(t) = \exp[(\frac{\lambda_2+\lambda_1}{\lambda_2-\lambda_1})t]$$

$$X \prod_{n=0}^{\infty} \exp[-t(1-\alpha)\alpha^n(\lambda_1+\lambda_2)/(\lambda_2-\lambda_1)][\pi_1 e^{t(1-\alpha)\alpha^n} + \pi_2 e^{-t(1-\alpha)\alpha^n}].$$

Since

$$\prod_{n=0}^{\infty} \exp[-t(1-\alpha)\alpha^n(\lambda_1+\lambda_2)/(\lambda_2-\lambda_1)] = \exp[-(\frac{\lambda_2+\lambda_1}{\lambda_2-\lambda_1})t] ,$$

$$m_x(t) = \prod_{n=0}^{\infty} [\pi_1 e^{t(1-\alpha)\alpha^n} + \pi_2 e^{-t(1-\alpha)\alpha^n}] \ . \tag{2.2.5}$$

The characteristic function of the distribution of x is $\Phi_x(t)$ defined by

$$\Phi_x(t) = m_x(it) \ ,$$

where $i^2 = -1$. Thus, from equation (2.2.5) ,

$$\Phi_x(t) = \prod_{n=0}^{\infty} [\pi_1 e^{it(1-\alpha)\alpha^n} + \pi_2 e^{-it(1-\alpha)\alpha^n}] \ . \tag{2.2.6}$$

SPECIAL CASES

Suppose $\pi_1 = \pi_2 = \frac{1}{2}$. Then (2.2.6) becomes

$$\Phi_x(t) = \prod_{n=0}^{\infty} \cos[t(1-\alpha)\alpha^n] \ . \tag{2.2.7}$$

Examples of the types of limiting distribution to which this characteristic function gives rise are given below for several values of α ranging from 0 to 1.

(i) $\alpha = 0$. When $\alpha = 0$, equation (2.2.7) becomes $\Phi_x(t) = \cos t$. The binomial distribution of the random variable x which assumes values ± 1 with equal probability has characteristic function

$$\tfrac{1}{2}e^{it} + \tfrac{1}{2}e^{-it} = \cos t \ .$$

Thus, the limiting distribution for $\alpha = 0$ is this binomial distribution.

(ii) $\alpha = \frac{1}{4}$. In this case, equation (2.2.7) becomes

$$\Phi_x(t) = \prod_{n=0}^{\infty} \cos(3t/2^{2n+2}) \tag{2.2.8}$$

$$= \prod_{n=0}^{\infty} \cos(3t/2^{2n+2}) \prod_{n=0}^{\infty} \cos(3t/2^{2n+1}) / \prod_{n=0}^{\infty} \cos(3t/2^{2n+1})$$

$$= \prod_{n=1}^{\infty} \cos(3t/2^{n}) \prod_{n=0}^{\infty} \sec(3t/2^{2n+1}) .$$

By a well known formula [see, for example, Kac [13], page 1, 2],

$$(2.2.9) \qquad \prod_{n=1}^{\infty} \cos(3t/2^{n}) = \frac{\sin 3t}{3t} .$$

Thus,

$$(2.2.10) \qquad \Phi_x(t) = \frac{\sin 3t}{3t} \prod_{n=0}^{\infty} \sec(3t/2^{2n+1}) .$$

Now, if we let

$$\Psi_o(t) = \frac{\sin 3t}{3t}$$

and

$$\Psi_k(t) = \frac{\sin 3t}{3t} \prod_{n=0}^{k-1} \sec(3t/2^{2n+1}) \qquad (k \geq 1) ,$$

we observe that

$$\Phi_x(t) = \lim_{k \to \infty} \Psi_k(t) ,$$

$$\Psi_1(t) = \frac{\sin 3t}{3t} . \sec \frac{3t}{2}$$

$$= \frac{\sin(3t/2)}{(3t/2)}$$

$$= \Psi_o(t/4) \cos(3t/4) ,$$

$$\Psi_2(t) = \frac{\sin(3t/2)}{(3t/2)} . \sec(3t/2^3)$$

$$= \frac{\sin(3t/2^3)}{(3t/2^3)} \;.\; \cos\,(3t/4)$$

$$= \Psi_1(t/4)\,\cos\,(3t/4)\;,$$

and

$$\Psi_3(t) = \frac{\sin(3t/2^3)}{(3t/2^3)} \;.\; \cos\,(3t/4) \;.\; \sec\,(3t/2^5)$$

$$= \frac{\sin(3t/2^5)}{(3t/2^5)} \;.\; \cos\,(3t/4) \;.\; \cos\,(3t/2^4)$$

$$= \Psi_2(t/4)\,\cos\,(3t/4)\;.$$

This suggests the induction hypothesis,

(2.2.11)
$$\Psi_k(t) = \Psi_{k-1}(t/4)\,\cos\,(3t/4)$$

$$= \frac{\sin(3t/2^{2k-1})}{(3t/2^{2k-1})} \prod_{n=1}^{k-1} \cos\,(3t/2^{2n})\;.$$

Under this hypothesis,

$$\Psi_{k+1}(t) = \frac{\sin\,3t}{3t} \prod_{n=0}^{k} \sec(3t/2^{2n+1})$$

$$= \Psi_k(t)\,\sec(3t/2^{2k+1})$$

$$= \frac{\sin\,(3t/2^{2k-1})}{(3t/2^{2k-1})} \prod_{n=1}^{k-1} \cos\,(3t/2^{2n}).\sec(3t/2^{2k+1})$$

$$= \frac{\sin\,(3t/2^{2k+1})}{(3t/2^{2k+1})} \prod_{n=1}^{k} \cos\,(3t/2^{2n})$$

$$= \Psi_k(t/4)\,\cos\,(3t/4)\;.$$

This is the same result as that given by equation (2.2.11) when k is replaced by $(k+1)$ and hence, by induction, the assertion,

$$\Psi_n(t) = \Psi_{n-1}(t/4)\cos(3t/4) \qquad (n=1,2,3,\ \dots\ ), \tag{2.2.12}$$

is true.

Let $g_k(x)$ be the probability density function corresponding to the characteristic function $\Psi_k(t)$, so that

$$g_k(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \Psi_k(t)e^{-itx}dt$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} \Psi_{k-1}(t/4)\cos(3t/4)e^{-itx}dt.$$

Then, since $\Psi_{k-1}(t/4)$ is an even function of t, we have by a result given in [14] (page 7),

$$g_k(x) = 2[g_{k-1}(4x-3) + g_{k-1}(4x+3)] \qquad (k \geq 1)\ . \tag{2.2.13}$$

If $g_o(x)$ is known, equations (2.2.13) determine a sequence of distributions which tend to the limiting distribution corresponding to $\alpha = \frac{1}{4}$. But

$$g_o(x) = \frac{1}{\pi}\int_{o}^{\infty} \frac{\sin 3t \cos tx}{3t}\ dt$$

and by a result given in [15] (page 60),

$$g_o(x) = \frac{1}{6} \qquad (-3 < x < 3)$$

$$= 0 \qquad (x < -3,\ x > 3).$$

Thus,

$$g_1(x) = \frac{1}{3} \qquad (-\tfrac{3}{2} < x < \tfrac{3}{2})$$

$$= 0 \qquad (x < -\tfrac{3}{2},\ x > \tfrac{3}{2}),$$

$$
\begin{aligned}
g_2(x) &= \tfrac{2}{3} && (-\tfrac{9}{8} < x < -\tfrac{3}{8},\ \tfrac{3}{8} < x < \tfrac{9}{8}) \\
&= 0 && (x < -\tfrac{9}{8},\ -\tfrac{3}{8} < x < \tfrac{3}{8},\ x > \tfrac{9}{8}), \\
g_3(x) &= \tfrac{4}{3} && (-\tfrac{33}{32} < x < -\tfrac{27}{32},\ -\tfrac{21}{32} < x < -\tfrac{15}{32}, \\
& && \tfrac{15}{32} < x < \tfrac{21}{32},\ \tfrac{27}{32} < x < \tfrac{33}{32}) \\
&= 0 && (x < -\tfrac{33}{32},\ -\tfrac{27}{32} < x < -\tfrac{21}{32}, \\
& && -\tfrac{15}{32} < x < \tfrac{15}{32},\ \tfrac{21}{32} < x < \tfrac{27}{32},\ x > \tfrac{33}{32}),
\end{aligned}
$$

and so forth.

Denote the 2^{n-1} intervals of nonzero probability density which occur at the $n^{\underline{th}}$ stage by $I_{n,m}(m=1,2,\ldots,2^{n-1};\ n \geq 1)$. The 2^n intervals of nonzero probability density which occur at stage $n+1$ are constructed from the intervals of the preceeding stage in the following way: divide each interval, $I_{n,m}$, into eight equal subintervals, say, $i_{n,m,1}$, $i_{n,m,2}$, ..., $i_{n,m,8}$ from left to right and retain as the two new intervals obtained from $I_{n,m}$, the union of $i_{n,m,2}$ and $i_{n,m,3}$ and the union of $i_{n,m,6}$ and $i_{n,m,7}$. The value of the probability density function over each of these new intervals is double that of the probability density function over the intervals of the n^{th} stage. From this construction, it follows that the length of each of the intervals at the n^{th} stage is $3/2^{2(n-1)}$ and the value of the probability density function over these intervals is $2^{n-1}/3$. The total length of the intervals at stage n is $3 \cdot 2^{n-1}/2^{2(n-1)}$ which approachs 0 as n approaches infinity. Thus, the asymptotic distribution is concentrated on a set of points of Lebesgue measure 0. This set is constructed in much the same

way as the Cantor Ternary set. This suggests that the limiting distribution is concentrated on a nonenumerable set of points of Lebesgue measure 0 and that the cumulative distribution function is not absolutely continuous.

(iii) $\alpha = \frac{1}{2}$. When $\alpha = \frac{1}{2}$, equation (2.2.7) becomes

$$\begin{aligned}\Phi_x(t) &= \prod_{n=1}^{\infty} \cos(t/2^n) \\ &= \sin t/t \ .\end{aligned}$$

Thus, the asymptotic distribution is uniform over (-1, 1).

(iv) $\alpha = 1/\sqrt{2}$. In this case,

$$\begin{aligned}\Phi_x(t) &= \prod_{n=0}^{\infty} \cos[(\sqrt{2}-1)t\ 2^{-(n+1)/2}] \\ &= \prod_{n=1}^{\infty} \cos[(\sqrt{2}-1)t\ 2^{-n}] \prod_{n=1}^{\infty} \cos[(2-\sqrt{2})t\ 2^{-n}] \\ &= \{\sin[(\sqrt{2}-1)t]/(\sqrt{2}-1)t\} \cdot \{\sin[(2-\sqrt{2})t]/(2-\sqrt{2})t\}\end{aligned}$$

Now, $\sin[(\sqrt{2}-1)t]/(\sqrt{2}-1)t$ and $\sin[(2-\sqrt{2})t]/(2-\sqrt{2})t$ are the characteristic functions of

$$\begin{aligned}f_{(x)} &= \frac{1}{2(\sqrt{2}-1)} && (-[\sqrt{2}-1] < x < [\sqrt{2}-1]) \\ &= 0 && (x < [\sqrt{2}-1]\ ,\ x > [\sqrt{2}-1])\end{aligned}$$

and

$$\begin{aligned}f_{(x)} &= \frac{1}{2\sqrt{2}(\sqrt{2}-1)} && (-\sqrt{2}[\sqrt{2}-1] < x < \sqrt{2}[\sqrt{2}-1]) \\ &= 0 && (x < -\sqrt{2}[\sqrt{2}-1]\ ,\ x > \sqrt{2}[\sqrt{2}-1]),\end{aligned}$$

respectively. Thus, $\Phi_x(t)$ is the characteristic function of the convolution of a uniform distribution over $(-[\sqrt{2}-1], [\sqrt{2}-1])$ with a

uniform distribution over $(-[2-\sqrt{2}]\ ,\ [2-\sqrt{2}])$.

In order to determine the limiting distribution, consider the more general case, where

$$(2.2.15)\qquad g_a(x) = \frac{1}{2a} \quad (-a < x < a)$$
$$= 0 \quad (x < -a,\ x > a)\ ,$$

$$(2.2.16)\qquad g_b(x) = \frac{1}{2b} \quad (-b < x < b)$$
$$= 0 \quad (x < -b,\ x > b)\ ,$$

and $a > b$. The convolution of $g_a(x)$ and $g_b(x)$ is

$$f(x) = \int_{-\infty}^{\infty} g_a(y)\, g_b(x-y)\, dy \ .$$

From (2.2.15) and (2.2.16), the product $g_a(y)\, g_b(x-y)$ is $\frac{1}{4ab}$ when y satisfies simultaneously the conditions, $-a < y < a$ and $x-b < y < x+b$; otherwise it is zero. Therefore,

$$f(x) = 0 \qquad (x + b < - a)\ ,$$

$$= \frac{1}{4ab}\int_{-a}^{x+b} dx = \frac{x+a+b}{4ab} \qquad (x-b < -a < x+b)\ ,$$

$$= \frac{1}{4ab}\int_{x-b}^{x+b} dx = \frac{1}{2a} \qquad (x-b > -a \ \text{ and } \ x+b < a)\ ,$$

$$= \frac{1}{4ab}\int_{x-b}^{a} dx = \frac{a+b-x}{4ab} \qquad (x-b < a < x+b)\ ,$$

$$= 0 \qquad (x-b > a)\ .$$

By applying these results to the case $a = 2-\sqrt{2}$ and $b = \sqrt{2}-1$ (so that $a+b = 1$ and $a-b = 3-2\sqrt{2}$), the probability density function, $f(x)$, of the limiting distribution is obtained. It is given by,

$$\begin{aligned} f(x) &= 1+x/[4\sqrt{2}(\sqrt{2}-1)^2] && (-1 < x < -[3-2\sqrt{2}]) , \\ &= 1/[2\sqrt{2}(\sqrt{2}-1)] && (-[3-2\sqrt{2}] < x < [3-2\sqrt{2}]) , \\ &= 1-x/[4\sqrt{2}(\sqrt{2}-1)^2] && ([3-2\sqrt{2}] < x < 1) , \\ &= 0 && (x < -1,\ x > 1) . \end{aligned}$$

(v) $\alpha = 1$. Here, $\Phi_x(t) = 1$ and so the asymptotic distribution is given by

$$\begin{aligned} &P[x=0] = 1 \\ &P[x=w] = 0 \qquad w \neq 0 . \end{aligned}$$

CHAPTER III

ADDITIONAL RESULTS FOR CERTAIN LIMITING DISTRIBUTIONS

Let $g(x ; a)$ denote the distribution function which is 0, $\frac{1}{2}$ or 1 according as $x < -a$, $-a \leq x < a$ or $x \geq a$. Then $g(x ; a)$ is called a symmetric Bernoulli distribution function and has as its characteristic function, $\cos(at)$. The distribution function, say, $\sigma(x)$, of the convolution of infinitely many such distributions is the distribution function of the quantity $x_1 + x_2 + \ldots$, where the x_i for $i = 1, 2, \ldots$ are independent random variables such that x_i is $-a_i$ or a_i with probability $\frac{1}{2}$. The characteristic function, say, $L(t ; \sigma)$, of $\sigma(x)$ is, therefore,

$$L(t ; \sigma) = \prod_{n=1}^{\infty} \cos(a_n t)$$

provided that this infinite product converges. The characteristic function given by equation (2.2.7) is a particular case of this function. Thus many properties of the desired limiting distributions are given in studies relating to $L(t ; \sigma)$. Some of these properties are stated below.

Jessen and Wintner [16] have shown that $\sigma(x)$ is either absolutely continuous or purely singular. Wintner [12] proved that when $a_n = 2^{-n/h}$ ($n = 1, 2, \ldots ; h = 1, 2, \ldots$), $\sigma(x)$ is absolutely continuous and acquires derivatives of arbitrarily high order for every x as $h \to \infty$, there being assured, for fixed h, the existence of a continuous $(h-1)^{th}$ derivative for every x. Kershner and Wintner [17] have shown that for the case $a_n = a^n (0 < a < 1)$, $\sigma(x)$ is singular and concentrated on a set of measure

zero when $a < \frac{1}{2}$, that $\sigma(x)$ is concentrated on the Cantor set when $a = \frac{1}{3}$, and that $\sigma(x) = \frac{x+1}{2}$ if $-1 \leq x \leq 1$ and $a = \frac{1}{2}$. On the basis of the work done by these authors, Erdos [18] proved that for the case $a_n = a^n$, there exist values of a, with $\frac{1}{2} < a < 1$, for which $\sigma(x)$ is purely singular. Examples are $a = \frac{1}{2}(\sqrt{5} - 1)$ and the positive root of the cubic equation $a^3 + a^2 - 1 = 0$ (This root is between $1/\sqrt{2}$ and $1/\sqrt[3]{2}$).

From the summary of known results which has been given above and in Chapter II, it is clear that the general character of the asymptotic distribution is known for $0 \leq \alpha \leq \frac{1}{2}$ in processes involving two experimenter-controlled events and the equal alpha condition. With the restriction $\pi_1 = \pi_2 = \frac{1}{2}$ imposed, an explicit solution has been determined for $\alpha = 1/\sqrt{2}$ and a method suggested for determining the asymptotic distributions corresponding to the members of the sequence $\alpha_n = 1/\sqrt[n]{2}$, $n = 1, 2, \ldots$. Although it is known that purely singular solutions occur for $\pi_1 = \pi_2 = \frac{1}{2}$ and certain values of α in the interval $(\frac{1}{2}, 1)$, the set of values in this interval for which the solution is singular is not known.

The present chapter of this thesis presents some new results and includes an extension of some of these results to the more general process discussed in Section Three of Chapter I. In Section One, a set of functional equations for the characteristic function of the limiting distribution are derived from equation (2.2.1). These, in turn, yield a set of functional equations for the limiting cumulative probability distribution. One member of this latter set proves useful in Section Two and is suitable for use in the numerical approximation of the limiting distribution. A system of n functional equations for the limiting cumulative probability distribution

function yields an explicit solution for the raw moments of the limiting distribution. In Section Two a simple procedure is developed for computing the limiting probability density function corresponding to given members of the sequence, $\alpha_n = 1/\sqrt[n]{2}$ $(n = 1,2, \ldots)$. The character of the several solutions exhibited in Section Two suggests that the sequence of limiting distributions corresponding to the members of the sequence $\alpha_n = 1/\sqrt[n]{2}$ $(n = 1, 2, \ldots)$ tends in the limit to the normal distribution. This is proven (without recourse to the Central Limit Theorem) in Section Three. Finally, in Section Four, several basic results are extended to the case of learning processes involving t experimenter-controlled results.

SECTION ONE

If $m(t)$ denotes the moment generating function of the limiting distribution of response probabilities in a learning process involving two experimenter-controlled events and the equal alpha condition then [by equation (2.2.2)] $m(t)$ satisfies the functional equation

$$m(t) = m(\alpha t)[\pi_1 e^{t(1-\alpha)\lambda_1} + \pi_2 e^{t(1-\alpha)\lambda_2}] \tag{3.1.1}$$

Denote by $m_x(t)$ and $\Phi_x(t)$, respectively, the moment generating and characteristic functions of the distribution, over $(-1, 1)$, of the random variable x given by

$$x = \left[\frac{\lambda_2 + \lambda_1}{\lambda_2 - \lambda_1}\right] - \left[\frac{2p}{\lambda_2 - \lambda_1}\right] \quad (\lambda_2 \neq \lambda_1),$$

where p is the random variable of the limiting distribution of response probabilities. Then $\Phi_x(t) = m_x(it)$ and, using result (2.2.4),

$$\Phi_x(t) = \exp\left[it\left(\frac{\lambda_2 + \lambda_1}{\lambda_2 - \lambda_1}\right)\right] m\left(\frac{-2it}{\lambda_2 - \lambda_1}\right)$$

(3.1.2)

$$= \exp\left[it\left(\frac{\lambda_2 + \lambda_1}{\lambda_2 - \lambda_1}\right)\right] \Phi\left(\frac{-2t}{\lambda_2 - \lambda_1}\right) ,$$

where $\Phi(t) = m(it)$ and, by equation (3.1.1), $\Phi(t)$ satisfies

(3.1.3) $$\Phi(t) = \Phi(\alpha t)[\pi_1 e^{it(1-\alpha)\lambda_1} + \pi_2 e^{it(1-\alpha)\lambda_2}].$$

Thus, from (3.1.2) and (3.1.3),

$$\Phi(t) = \Phi_x[-\frac{t}{2}(\lambda_2 - \lambda_1)] \exp[\frac{it}{2}(\lambda_2 + \lambda_1)]$$

and

$$\Phi_x[-\frac{t}{2}(\lambda_2 - \lambda_1)] = \exp[-\frac{it(1-\alpha)}{2}(\lambda_2 + \lambda_1)] \Phi_x[-\frac{\alpha t}{2}(\lambda_2 - \lambda_1)]$$
$$\times [\pi_1 e^{it(1-\alpha)\lambda_1} + \pi_2 e^{it(1-\alpha)\lambda_2}]$$

or

$$\Phi_x(t) = \Phi_x(\alpha t) \exp[it(1-\alpha)(\frac{\lambda_2 + \lambda_1}{\lambda_2 - \lambda_1})]$$
$$\times [\pi_1 e^{-2it(1-\alpha)\lambda_1/(\lambda_2-\lambda_1)} + \pi_2 e^{-2it(1-\alpha)\lambda_2/(\lambda_2-\lambda_1)}]$$

giving

(3.1.4) $$\Phi_x(t) = \Phi_x(\alpha t)[\pi_1 e^{it(1-\alpha)} + \pi_2 e^{-it(1-\alpha)}] .$$

Equation (3.1.4) implies that

$$\Phi_x(t) = \Phi_x(\alpha^2 t) \prod_{n=1}^{2} [\pi_1 e^{it(1-\alpha)\alpha^{n-1}} + \pi_2 e^{-it(1-\alpha)\alpha^{n-1}}]$$

and

$$\Phi_x(t) = \Phi_x(\alpha^3 t) \prod_{n=1}^{3} [\pi_1 e^{it(1-\alpha)\alpha^{n-1}} + \pi_2 e^{-it(1-\alpha)\alpha^{n-1}}] .$$

Continuing in this way, we obtain

$$\Phi_x(t) = \Phi_x(\alpha^r t) \prod_{n=1}^{r} [\pi_1 e^{it(1-\alpha)\alpha^{n-1}} + \pi_2 e^{-it(1-\alpha)\alpha^{n-1}}], (r \geq 1). \tag{3.1.5}$$

In the expansion of

$$\prod_{n=1}^{r} [\pi_1 e^{it(1-\alpha)\alpha^{n-1}} + \pi_2 e^{-it(1-\alpha)\alpha^{n-1}}] ,$$

there are 2^r terms each of the form

$$\pi_1^m \ \pi_2^{r-m} \ \exp[-it(1-\alpha)(c_1 + c_2\alpha + \ldots + c_r\alpha^{r-1})] ,$$

where $c_j (j = 1, 2, \ldots , r)$ is ± 1 and m is the number of the c_j which have value -1. Let $X_{ri} (i = 1, 2, \ldots , 2^r)$ denote the possible values of

$$(1-\alpha)(c_1 + c_2\alpha + \ldots + c_r\alpha^{r-1}) \tag{3.1.6}$$

and let $m(X_{ri})$ be the number of the c_j which have value -1 when X_{ri} is expressed in the form of (3.1.6). Then,

$$\Phi_x(t) = \Phi_x(\alpha^r t) \sum_{j=1}^{2^r} \pi_1^{m(X_{rj})} \ \pi_2^{r-m(X_{rj})} \ e^{-itX_{rj}} . \tag{3.1.7}$$

By the inversion theorem (see, for example, Kendall [19] page 94), the cumulative probability distribution function, say, $F(x)$, corresponding to $\Phi_x(t)$ is given by

$$F(x) = F(0) + \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1 - e^{-ixt}}{it} \Phi_x(t)\, dt , \tag{3.1.8}$$

where $F(x) = 0$ for $x \leq -1$ and $F(x) = 1$ for $x \geq 1$, since x is distributed on $(-1, 1)$. Then, if a is any constant

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1 - e^{-ixt}}{it} \cdot e^{-ita} \cdot \Phi_x(\alpha^r t)\, dt$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1 - e^{-it(x+a)/\alpha^r}}{it} \cdot \Phi_x(t)\, dt - \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1-e^{-ita/\alpha^r}}{it} \cdot \Phi_x(t) dt,$$

so that, by equation (3.1.8),

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1-e^{-ixt}}{it} \cdot e^{-ita} \Phi_x(\alpha^r t)\, dt = F\left(\frac{x+a}{\alpha^r}\right) - F\left(\frac{a}{\alpha^r}\right) \qquad (r \geq 1) . \tag{3.1.9}$$

From equations (3.1.7) and (3.1.8) it follows that

$$F(x) - F(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1-e^{-ixt}}{it} \cdot \Phi_x(t)\, dt$$

$$= \sum_{j=1}^{2^r} \pi_1^{m(X_{rj})} \pi_2^{r-m(X_{rj})} \left\{ \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1-e^{-ixt}}{it} \cdot e^{-itX_{rj}} \cdot \Phi_x(\alpha^r t)\, dt \right\}$$

and hence, by equation (3.1.9), that

$$(3.1.10)\quad F(x) - F(0) = \sum_{i=1}^{2^r} \pi_1^{m(X_{ri})} \pi_2^{r-m(X_{ri})} \left[F\left(\frac{x+X_{ri}}{\alpha^r}\right) - F\left(\frac{X_{ri}}{\alpha^r}\right) \right] .$$

In particular,

$$F(-1) - F(0) = \sum_{i=1}^{2^r} \pi_1^{m(X_{ri})} \pi_2^{r-m(X_{ri})} \left[F\left(\frac{-1 + X_{ri}}{\alpha^r}\right) - F\left(\frac{X_{ri}}{\alpha^r}\right) \right] .$$

But,

$$X_{ri} \leq (1-\alpha)(1+\alpha+ \ldots + \alpha^{r-1}) = 1-\alpha^r$$

and hence

$$F\left(\frac{-1 + X_{ri}}{\alpha^r}\right) \leq F\left(\frac{-1 + 1 - \alpha^r}{\alpha^r}\right)$$

$$= F(-1) = 0$$

and since $F\left(\frac{-1 + X_{ri}}{\alpha^r}\right) \geq 0$, we must have $F\left(\frac{-1 + X_{ri}}{\alpha^r}\right) = 0 \quad (i = 1, 2, \ldots, 2^r)$.

Thus,

$$F(0) = \sum_{i=1}^{2^r} \pi_1^{m(X_{ri})} \pi_2^{r-m(X_{ri})} F\left(\frac{X_{ri}}{\alpha^r}\right)$$

and

$$(3.1.11)\quad F(x) = \sum_{i=1}^{2^r} \pi_1^{m(X_{ri})} \pi_2^{r-m(X_{ri})} F\left(\frac{x + X_{ri}}{\alpha^r}\right) \quad (r \geq 1) .$$

When $r = 1$, equation (3.1.11) becomes

$$(3.1.12)\quad F(x) = \pi_2 F\left(\frac{x + 1 - \alpha}{\alpha}\right) + \pi_1 F\left(\frac{x - 1 + \alpha}{\alpha}\right) .$$

This result will be applied in Section Two.

Equations (3.1.11) imply that

$$(3.1.13) \qquad d\,F(x) = \sum_{i=1}^{2^r} \pi_1^{m(X_{ri})} \pi_2^{r-m(X_{ri})} d\,F\left(\frac{x + X_{ri}}{\alpha^r}\right) \quad .$$

Let V_s denote the s^{th} raw moment of the distribution of the random variable, x . Then, since $d\,F(x) = 0$ when $x \leq -1$ and $x \geq 1$, we find from equation (3.1.13) that

$$V_s = \int_{-1}^{1} x^s dF(x) = \sum_{i=1}^{2^r} \pi_1^{m(X_{ri})} \pi_2^{r-m(X_{ri})} \int_{-\alpha^r - X_{ri}}^{\alpha^r - X_{ri}} x^s dF\left(\frac{x + X_{ri}}{\alpha^r}\right)$$

$$= \sum_{i=1}^{2^r} \pi_1^{m(X_{ri})} \pi_2^{r-m(X_{ri})} \int_{-1}^{1} [\alpha^r x - X_{ri}]^s \, dF(x)$$

$$= \sum_{i=1}^{2^r} \pi_1^{m(X_{ri})} \pi_2^{r-m(X_{ri})} \left\{ \sum_{k=0}^{s} \binom{s}{k} \alpha^{kr} (-X_{ri})^{s-k} V_k \right\}$$

$$= \sum_{k=0}^{s} \binom{s}{k} \alpha^{kr} \, V_k \, S_r^{(s-k)} \qquad (r \geq 1, \; s \geq 1) \; ,$$

where

$$S_r^{(s-k)} = \sum_{i=1}^{2^r} \pi_1^{m(X_{ri})} \pi_2^{r-m(X_{ri})} (-X_{ri})^{s-k} \; .$$

Thus,

$$(3.1.14) \qquad V_s(1-\alpha^{rs}) = \sum_{k=0}^{s-1} \binom{s}{k} \alpha^{kr} \, V_k \, S_r^{(s-k)} \quad (r \geq 1, \; s \geq 1)$$

$$V_0 = 1 \quad .$$

From equations (3.1.14) we obtain a system of s equations in the unknown quantities $V_1, V_2, \ldots, V_s$ by choosing the equations corresponding to $r = 1, 2, \ldots, s$. That is,

$$V_s(1-\alpha^s) - \sum_{k=1}^{s-1} \binom{s}{s-k} \alpha^{s-k} V_{s-k} S_1^{(k)} = S_1^{(s)}$$

$$V_s(1-\alpha^{2s}) - \sum_{k=1}^{s-1} \binom{s}{s-k} \alpha^{2(s-k)} V_{s-k} S_2^{(k)} = S_2^{(s)} \tag{3.1.15}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$V_s(1-\alpha^{s^2}) - \sum_{k=1}^{s-1} \binom{s}{s-k} \alpha^{s(s-k)} V_{s-k} S_s^{(k)} = S_s^{(s)} \qquad (s \geq 1) .$$

When $\pi_1 = \pi_2 = \frac{1}{2}$,

$$S_r^{(s-k)} = (\tfrac{1}{2})^r \sum_{i=1}^{2^r} (-X_{ri})^{s-k}$$

is zero if $s-k$ is odd since the terms in the sum will then cancel in pairs. At the same time $V_{2S+1} = 0$ $(S = 0, 1, 2, \ldots)$ since the characteristic function of $F(x)$ [see equation (2.2.7)] is even. Thus, for the case $\pi_1 = \pi_2 = \frac{1}{2}$ we obtain, from equations (3.1.15), the following system of equations in the unknown quantities $V_{2s}, V_{2s-2}, \ldots, V_2$:

$$\begin{aligned}
&V_{2s}(1-\alpha^{2s}) - \sum_{k=1}^{s-1} \binom{2s}{2s-2k} \alpha^{2(s-k)} V_{2s-2k}\, S_1^{(2k)} = S_1^{(2s)} \\
&V_{2s}(1-\alpha^{4s}) - \sum_{k=1}^{s-1} \binom{2s}{2s-2k} \alpha^{2(2s-2k)} V_{2s-2k}\, S_2^{(2k)} = S_2^{(2s)} \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&V_{2s}(1-\alpha^{2s^2}) - \sum_{k=1}^{s-1} \binom{2s}{2s-2k} \alpha^{s(2s-2k)} V_{2s-2k} S_s^{(2k)} = S_s^{(2s)} \quad (s \geq 1).
\end{aligned} \tag{3.1.16}$$

Provided that $\pi_1 \neq \pi_2$, explicit solutions are obtained for V_s $(s = 1, 2, \ldots)$ by solving equations (3.1.15). The determinant of the coefficients of $V_s, V_{s-1}, \ldots,$ and V_1, say, D_{s,π_1}, is given by

$$D_{s,\pi_1} = \left[\prod_{k=1}^{s-1} \binom{s}{s-k}\right] \left| a_{ij,\pi_1} \right|$$

where

$$a_{i1,\pi_1} = (1-\alpha^{is}) \qquad (i = 1, 2, \ldots, s)$$

and

$$a_{ij,\pi_1} = -\alpha^{i(s-j+1)}\, S_i^{(j-1)} \qquad (i = 1, 2, \ldots, s;\ j = 2, 3, \ldots, s)$$

Thus,

$$\begin{aligned}
V_s &= \left[\prod_{k=1}^{s-1} \binom{s}{s-k}\right] \left| b_{ij,\pi_1} \right| / D_{s,\pi_1} \\
&= \left| b_{ij,\pi_1} \right| / \left| a_{ij,\pi_1} \right| ,
\end{aligned} \tag{3.1.17}$$

where

$$b_{i1,\pi_1} = S_i^{(s)} \qquad (i = 1, 2, \ldots, s)$$

and

$$b_{ij,\pi_1} = a_{ij,\pi_1} \quad (i = 1, 2, \ldots, s \; ; \; j = 2, 3, \ldots, s).$$

On the other hand, if $\pi_1 = \pi_2 = \frac{1}{2}$, V_{2s} $(s = 1, 2, \ldots)$ is found from equations (3.1.16). If $D_{2s,\frac{1}{2}}$ is the determinant of the coefficient of $V_2, V_4, \ldots, V_{2s}$,

$$D_{2s,\frac{1}{2}} = \left[\prod_{k=1}^{s-1} \binom{2s}{2s-2k}\right] \left| a_{ij,\frac{1}{2}} \right|$$

where

$$a_{i1,\frac{1}{2}} = (1-\alpha^{2is}) \qquad (i=1, 2, \ldots, s)$$

and

$$a_{ij,\frac{1}{2}} = -\alpha^{i(2s-2j+2)} S_i^{(2j-2)} \qquad (i=1,2, \ldots, s; \; j=2,3,\ldots,s).$$

Thus, for $\pi_1 = \pi_2 = \frac{1}{2}$,

$$V_{2s} = \left[\prod_{k=1}^{s-1} \binom{2s}{2s-2k}\right] \left| b_{ij,\frac{1}{2}} \right| / D_{2s,\frac{1}{2}} \tag{3.1.18}$$

$$= \left| b_{ij,\frac{1}{2}} \right| / \left| a_{ij,\frac{1}{2}} \right|,$$

where

$$b_{i1,\frac{1}{2}} = S_i^{(2s)} \qquad (i = 1, 2, \ldots, s)$$

and

$$b_{ij,\frac{1}{2}} = a_{ij,\frac{1}{2}} \qquad (i = 1,2,\ldots,s\ ;\ j = 2,3,\ldots,s).$$

We observe that when $\pi_1 = \pi_2 = \frac{1}{2}$, the first of equations (3.1.16) gives

$$V_2 = (1- \alpha^2)^{-1} \ s_1^{(2)} \quad .$$

Since $X_{11} = -(1-\alpha)$ and $X_{12} = (1-\alpha)$,

$$\begin{aligned} s_1^{(2)} &= \tfrac{1}{2}[(1-\alpha)^2 + (1-\alpha)^2] \\ &= (1-\alpha)^2 \end{aligned}$$

and

(3.1.19) $$V_2 = \frac{1 - \alpha}{1 + \alpha} \quad .$$

SECTION TWO

Consider a learning process with two experimenter controlled events, E_1 and E_2 for which $\alpha_1 = \alpha_2 = 1/\sqrt[n]{2}$, where n is some positive integer. If, in addition, $\pi_1 = \pi_2 = \frac{1}{2}$, $\Phi_x(t)$ satisfies, by equation (2.2.7),

(3.2.1) $$\Phi_x(t) = \prod_{j=0}^{\infty} \cos\ [t(1-1/\sqrt[n]{2})(\sqrt[n]{2})^{-j}] \ ,$$

where $\Phi_x(t)$ is the characteristic function of the distribution of the random variable

$$x = \left[\frac{\lambda_2 + \lambda_1}{\lambda_2 - \lambda_1}\right] - \left[\frac{2p}{\lambda_2-\lambda_1}\right] , \quad (\lambda_2 \neq \lambda_1) \ ,$$

and p is the random variable of the limiting distribution of response probabilities for the process just described.

Since, by formula (2.2.9) ,

$$\prod_{j=1}^{\infty} \cos\left[\frac{2t(1-1/\sqrt[n]{2})2^{-(k-1)/n}}{2^j}\right] = \frac{\sin[2t(1-1/\sqrt[n]{2})2^{-(k-1)/n}]}{[2t(1-1/\sqrt[n]{2})\,2^{-(k-1)/n}]} ,$$

we have, from equation (3.2.1),

$$\Phi_x(t) = \prod_{k=1}^{n} \left\{\prod_{j=1}^{\infty} \cos\left[\frac{2t(1-1/\sqrt[n]{2})2^{-(k-1)/n}}{2^j}\right]\right\}$$

$$= \prod_{k=1}^{n} \frac{\sin[2t(1-1/\sqrt[n]{2})2^{-(k-1)/n}]}{[2t(1-1/\sqrt[n]{2})2^{-(k-1)/n}]}$$

or

$$(3.2.2) \qquad \Phi_x(t) = (C_n'/t^n) \prod_{k=1}^{n} \sin(b_k t) ,$$

where

$$C_n' = 1/\left[(\sqrt[n]{2} - 1)^n \cdot 2^{(n-1)/2}\right]$$

and

$$b_k = 2(\sqrt[n]{2} - 1)2^{-k/n} .$$

We observe that if $F_k(x)$ is the uniform distribution over $(-b_k, b_k)$, $\Phi_x(t)$ is the characteristic function of the convolution of $F_1(x), F_2(x), \ldots ,$ and $F_n(x)$.

Kemeny and Snell [9] page 229 have shown that in the case being considered, the limiting probability density function, say, $f_n(x)$, exists for all x. It is given by

$$f_n(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ixt} \Phi_x(t)\, dt \; . \tag{3.2.3}$$

Now, in the expansion of

$$(2i)^n \prod_{k=1}^{n} \sin(h_k t)$$

$$= \prod_{k=1}^{n} \left[\exp[2it(\sqrt[n]{2}-1)2^{-k/n}] - \exp[-2it(\sqrt[n]{2}-1)2^{-k/n}] \right] ,$$

there are 2^n terms, each of which may be written as

$$(-1)^m \exp\{it(\sqrt[n]{2}-1)[c_1+c_2(\sqrt[n]{2}) + \ldots + c_n(\sqrt[n]{2})^{n-1}]\} ,$$

where c_i $(i = 1, 2, \ldots, n)$ is ± 1 and m is the number of the c_i which have value -1. Let X_{ni} $(i = 1, 2, \ldots, 2^n)$, with $X_{ni} < X_{n,i+1}$, denote the possible values of

$$(\sqrt[n]{2}-1)[c_1+c_2(\sqrt[n]{2}) + \ldots + c_n(\sqrt[n]{2})^{n-1}] \tag{3.2.4}$$

and $m(X_{ni})$ be the number of the c_i which have value -1 when X_{ni} is expressed in the form of (3.2.4). Then

$$\Phi_x(t) = \left[C_n'/(2it)^n \right] \sum_{j=1}^{2^n} (-1)^{m(X_{nj})} e^{itX_{nj}} , \tag{3.2.5}$$

and using (3.2.3),

$$f_n(x) = \frac{C_n'}{2^{n+1}\pi} \int_{-\infty}^{\infty} \left\{ \sum_{j=1}^{2^n} \frac{(-1)^{m(X_{nj})}\, e^{-it(x-X_{nj})}}{(it)^n} \right\} dt \; . \tag{3.2.6}$$

It will now be shown that

$$(3.2.7) \qquad \frac{d^r f_n(x)}{dx^r} = \frac{C'_n}{2^{n+1}\pi} \int_{-\infty}^{\infty} \frac{d^r}{dx^r}\left[\sum_{j=1}^{2^n} \frac{(-1)^{m(X_{nj})} e^{-it(x-X_{nj})}}{(it)^n}\right] dt$$

for $r \geq 1$.

The integrand on the right hand side of equation (3.2.7) is

$$\frac{d^r}{dx^r}\left[\sum_{j=1}^{2^n} \frac{(-1)^{m(X_{nj})} e^{-it(x-X_{nj})}}{(it)^n}\right]$$

$$= \sum_{j=1}^{2^n} \frac{(-1)^{r+m(X_{nj})} e^{-it(x-X_{nj})}}{(it)^{n-r}}$$

$$= \frac{(-1)^r 2^n i^r}{t^{n-r}} \cdot e^{-itx} \prod_{k=1}^{n} \sin(b_k t)$$

which tends to 0 as t → 0 and is continuous in t for all x . The corresponding integral (see, for example, Pierpont [20] page 453) exists and is uniformly convergent for all x (Ibid., page 467). Assertion (3.2.7) follows (Ibid., page 493) from these observations.

When $r = n-1$, equation (3.2.7) becomes

$$(3.2.8) \qquad \frac{d^{n-1} f_n(x)}{dx^{n-1}} = \frac{(-1)^{n-1} C'_n}{2^{n+1}\pi} \int_{-\infty}^{\infty} \left[\sum_{j=1}^{2^n} \frac{(-1)^{m(X_{nj})}}{it} \frac{e^{-it(x-X_{nj})}}{}\right] dt$$

$$= \frac{(-1)^n C'_n}{2^n \pi} \sum_{i=1}^{2^n} (-1)^{m(X_{ni})} \int_0^{\infty} \frac{\sin[t(x-X_{ni})]}{t}\, dt \ .$$

But (see, for example, Kendall [19] page 95),

$$(3.2.9) \qquad \int_0^{\infty} \frac{\sin[t(x-X_{ni})]}{t}\,dt = \begin{cases} -\frac{\pi}{2} & x < X_{ni} \\ 0 & x = X_{ni} \\ \frac{\pi}{2} & x > X_{ni} \end{cases}$$

for $i = 1, 2, \ldots, 2^n$. Thus $\frac{d^{n-1} f_n(x)}{dx^{n-1}}$ is constant for all x and $f_n(x)$ is piecewise a polynomial of degree not exceeding $(n-1)$.

Now $X_{n1} = -(\sqrt[n]{2}-1)[1 + \sqrt[n]{2} + \ldots + (\sqrt[n]{2})^{n-1}] = -1$ and $X_{n2} = -(\sqrt[n]{2}-1)[-1 + \sqrt[n]{2} + \ldots + (\sqrt[n]{2})^{n-1}] = -3 + 2(\sqrt[n]{2})$. When $-1 < x < -3 + 2(\sqrt[n]{2})$,

$$\frac{d^{n-1}}{dx^{n-1}}[f_n(x)] = \frac{(-1)^n C_n'}{2^n \pi}\left\{(-1)^n \frac{\pi}{2} - \frac{\pi}{2}\sum_{i=2}^{2^n} (-1)^{m(X_{ni})}\right\}$$

$$= \frac{(-1)^r C_n'}{2^n \pi}\left\{(-1)^n \pi - \frac{\pi}{2}\left[\sum_{r=0}^{n} \binom{n}{r}(-1)^r\right]\right\}$$

$$= C_n'/2^n .$$

Thus, when $-1 < x < -3 + 2(\sqrt[n]{2})$ the coefficient of x^{n-1} in $f_n(x)$ is

$$(3.2.10) \qquad \frac{C_n'}{(n-1)!\,2^n} = 1/\left[(n-1)!\ (\sqrt[n]{2}-1)^n\, 2^{(3n-1)/2}\right].$$

It was shown [see equation (3.1.12)] that the cumulative probability distribution function, $F(x)$, of

$$x = \frac{\lambda_2 + \lambda_1}{\lambda_2 - \lambda_1} - \frac{2p}{\lambda_2 - \lambda_1} \qquad (\lambda_2 \neq \lambda_1)$$

satisfies, for general α and π_1,

$$\begin{aligned} F(x) &= \pi_2 F\left(\frac{x+1-\alpha}{\alpha}\right) + \pi_1 F\left(\frac{x-1+\alpha}{\alpha}\right) && (-\infty < x < \infty) \\ (3.2.11) \qquad &= 0 && (x \leq -1) \\ &= 1 && (x \geq 1) . \end{aligned}$$

When $\pi_1 = \pi_2 = \frac{1}{2}$ and $\alpha = 1/\sqrt[n]{2}$ $(n \geq 1)$, equation (3.2.11) becomes

$$\begin{aligned} 2F(x) &= F[(\sqrt[n]{2})x + \sqrt[n]{2} - 1] + F[(\sqrt[n]{2})x - \sqrt[n]{2} + 1] && (-\infty < x < \infty) , \\ &= 0 && (x \leq 1) , \\ &= 2 && (x \geq 1) . \end{aligned}$$

Thus, the probability density function $f_n(x)$ satisfies

$$\begin{aligned} (3.2.12) \quad 2f_n(x) &= \sqrt[n]{2}\{f_n[(\sqrt[n]{2})x + \sqrt[n]{2} - 1] + f_n[(\sqrt[n]{2})x - \sqrt[n]{2} + 1]\} && (-\infty < x < \infty) , \\ &= 0 && (x \leq -1,\ x \geq 1) , \end{aligned}$$

or

$$\begin{aligned} (3.2.13) \quad f_n(x) &= 2^{(n-1)/n} f_n\left(\frac{x - \sqrt[n]{2} + 1}{\sqrt[n]{2}}\right) - f_n[x + 2(1 - \sqrt[n]{2})] && (-\infty < x < \infty) , \\ &= 0 && (x \leq -1,\ x \geq 1) . \end{aligned}$$

Now, when

$$-1 < x < -3 + 2(\sqrt[n]{2}) , \quad x + 2(1 - \sqrt[n]{2}) < -1$$

and $f_n[x + 2(1 - \sqrt[n]{2})] = 0$. Thus, for $-1 < x < -3 + 2(\sqrt[n]{2})$,

$f_n(x)$ satisfies

$$(3.2.14) \qquad f_n(x) = 2^{(n-1)/n} \; f_n\left(\frac{x - \sqrt[n]{2} + 1}{\sqrt[n]{2}}\right) .$$

This equation has as its solution

$$(3.2.15) \qquad f_n(x) = C_n(1 + x)^{n-1} ,$$

where $$C_n = 1/\left[(n-1)! \; (\sqrt[n]{2}-1)^n \, 2^{(3n-1)/2}\right]$$

in order that the coefficient of x^{n-1} be the same as that given in equation (3.2.10).

For convenience denote by $T_{n1}(x)$ and $T_{n2}(x)$, respectively, the transformations, $\dfrac{x - \sqrt[n]{2} + 1}{\sqrt[n]{2}}$ and $x + 2(1 - \sqrt[n]{2})$. For all x such that

$-1 < T_{n1}(x) < -3 + 2(\sqrt[n]{2})$ and $-1 < T_{n2}(x) < -3 + 2(\sqrt[n]{2})$, that is, for

$-3 + 2(\sqrt[n]{2}) < x < -1 - 2(\sqrt[n]{2}) + 2(\sqrt[n]{2})^2$,

$$f_n(x) = C_n\{(x+1)^{n-1} - [x + 3 - 2(\sqrt[n]{2})]^{n-1}\} .$$

When $T_{n2}(x)$ is in the interval $(-1, -3 + 2(\sqrt[n]{2}))$ and $T_{n1}(x)$ is in $(-3 + 2(\sqrt[n]{2}), -1 - 2(\sqrt[n]{2}) + 2(\sqrt[n]{2})^2)$,

$$(3.2.16) \qquad f_n(x) = C_n\{(x+1)^{n-1} - [x+3-2(\sqrt[n]{2})]^{n-1} - [x+1+2(\sqrt[n]{2})-2(\sqrt[n]{2})^2]^{n-1}\} .$$

The upper boundary for the set of values of x for which (3.2.16) holds is the minimum of x_1 and x_2 found by solving

$$T_{n1}(x_1) = -1 - 2(\sqrt[n]{2}) + 2(\sqrt[n]{2})^2 \quad \text{and}$$

$$T_{n2}(x_2) = -3 + 2(\sqrt[n]{2}) .$$

While it appears that this procedure could be continued and $f_n(x)$ determined, the problem of determining the remaining boundaries sequentially, in this manner, has not been solved for general n. Equation (3.2.8) suggests that these boundaries are $X_{n1}, X_{n2}, \ldots, X_{n,2^n}$, which are found when the possible values of

$$(3.2.17) \qquad (\sqrt[n]{2}-1)[c_1 + c_2(\sqrt[n]{2}) + \ldots + c_n(\sqrt[n]{2})^{n-1}] ,$$

where c_i $(i = 1, 2, \ldots, n)$ is ± 1, are ranked in order of magnitude. If $f_n(x)$ is known for $T_{n1}(X_{n,i-1}) < x < T_{n1}(X_{ni})$ and $T_{n2}(X_{n,i-1}) < x < T_{n2}(X_{ni})$, equation (3.2.13) gives $f_n(x)$ for $X_{n,i-1} < x < X_{ni}$. Since $f_n(x)$ is known [see equation (3.2.15)] for $-1 = X_{n1} < x < X_{n2} = -3 + 2(\sqrt[n]{2})$, equation (3.2.13) can be applied repeatedly to give an explicit solution for $f_n(x)$ $(-1 < x < 1)$ provided that the values of X_{ni}, $T_{n1}(X_{ni})$, and $T_{n2}(X_{ni})$ are known for $i = 1, 2, \ldots, 2^n$.

However, it has not been possible to determine the $X_{ni}(i=1,2, \ldots, 2^n)$ for general n. For a given value of n, they may be found by using, as a guide, decimal approximations of the quantities given by (3.1.17). From these results, $T_{n1}(X_{ni})$ and $T_{n2}(X_{ni})$ $(i = 1, 2, \ldots, 2^n)$ together with their decimal approximations are easily determined. With the aid of these approximations, $f_n(x)$ is then found by means of equation (3.2.13). The method is illustrated, in the follwoing discussion, by the calculation of $f_3(x)$, $f_4(x)$, and $f_5(x)$ [Since $f_n(x)$ is a symmetric function these computations are carried out for $-1 < x < X_{n,2^{n-1}+1}$ $(n = 3, 4, 5)$]. Decimal approximations for X_{ni}, $T_{n1}(X_{ni})$ and $T_{n2}(X_{ni})$ will be denoted by A_{ni}, $D_{n1,i}$ and $D_{n2,i}$, respectively $(n = 3, 4, 5; i = 1, 2, \ldots, 2^{n-1}+1)$.

(i) $n = 3$. In this case $\alpha = 1/\sqrt[3]{2}$ and equation (3.2.13) becomes

$$(3.2.18) \qquad f_3(x) = 2^{2/3} f_3\left(\frac{x-\sqrt[3]{2}+1}{\sqrt[3]{2}}\right) - f_3[x+2(1-\sqrt[3]{2})] \quad (-\infty < x < \infty) ,$$

$$= 0 \qquad (x \leq -1,\ x \geq 1) .$$

By equation (3.2.15),

$$(3.2.19) \qquad f_3(x) = \frac{(x+1)^2}{32(\sqrt[3]{2}-1)^3} \qquad (-1 < x < -3 + 2(\sqrt[3]{2})) .$$

Table 3.1 gives some preliminary calculations which are necessary in the determination of $f_3(x)$ by the method discussed above.

i	X_{3i}	A_{3i}	$D_{31,i}$	$D_{32,i}$
1	-1	-1	-1	
2	$-3+2(\sqrt[3]{2})$	-0.480	-0.587	-1
3	$-1-2(\sqrt[3]{2})+2(\sqrt[3]{2})^2$	-0.345	-0.480	-0.865
4	$3-2(\sqrt[3]{2})^2$	-0.175	-0.345	-0.695
5	$-3+2(\sqrt[3]{2})^2$	$+0.175$	-0.068	-0.345

Table 3.1 Values of X_{3i} together with the decimal approximations, A_{3i}, $D_{31,i}$ and $D_{32,i}$ computed for X_{3i}, $T_{31}(X_{3i})$ and $T_{32}(X_{3i})$, respectively ($i = 1,2,3,4,5$).

Using equations (3.2.18) and (3.2.19) and the results given in Table 3.1, it follows that

$$f_3(x) = (x+1)^2/32(\sqrt[3]{2}-1)^3 \qquad (-1 < x < -3 + 2(\sqrt[3]{2})) ,$$

$$= \frac{1}{32(\sqrt[3]{2}-1)^3}\{[x+1]^2-[x+3-2(\sqrt[3]{2})]^2\}$$

(3.2.20)
$$(-3+2(\sqrt[3]{2}) < x < -1 -2(\sqrt[3]{2})+2(\sqrt[3]{2})^2) ,$$

$$= \frac{1}{32(\sqrt[3]{2}-1)^3}\{[x+1]^2-[x+3-2(\sqrt[3]{2})]^2-[x+1+2(\sqrt[3]{2})-2(\sqrt[3]{2})^2]^2\}$$

$$(-1-2(\sqrt[3]{2})+2(\sqrt[3]{2})^2 < x < 3 -2(\sqrt[3]{2})^2)$$

$$= \frac{1}{32(\sqrt[3]{2}-1)^3}\{[x+1]^2-[x+3-2(\sqrt[3]{2})]^2-[x+1+2(\sqrt[3]{2})-2(\sqrt[3]{2})^2]^2 -[x-3+2(\sqrt[3]{2})^2]^2\}$$

$$(3-2(\sqrt[3]{2})^2 < x < -3 + 2(\sqrt[3]{2})^2) .$$

(ii) $n = 4$. Here $\alpha = 1/\sqrt[4]{2}$. Equations (3.2.13) and (3.2.15) become

(3.2.21)
$$f_4(x) = 2^{3/4}f_4\left(\frac{x-2^{1/4}+1}{2^{1/4}}\right) - f_4[x+2(1-2^{1/4})] \qquad (-\infty< x <\infty)$$

$$= 0 \qquad (x \leq -1,\ x \geq 1)$$

and

(3.2.22)
$$f_4(x) = c_4(x+1)^3 ,$$

respectively, where

$$c_4 = 1/\left[6(2^{1/4}-1)^4\ 2^{11/2}\right].$$

Table 3.2 is the counterpart of Table 3.1 and contains results necessary in the calculation of $f_4(x)$.

i	X_{4i}	A_{4i}	$D_{41,i}$	$D_{42,i}$
1	-1	-1	-1	
2	$-3+2(2^{1/4})$	-0.622	-0.682	-1
3	$-1-2(2^{1/4})+2(2^{1/2})$	-0.550	-0.622	-0.928
4	$-1-2(2^{1/2})+2(2^{3/4})$	-0.465	-0.550	-0.842
5	$3-2(2^{3/4})$	-0.364	-0.465	-0.742
6	$-3+2(2^{1/2})$	-0.172	-0.304	-0.550
7	$-3+2(2^{1/4})-2(2^{1/2})+2(2^{3/4})$	-0.086	-0.232	-0.465
8	$-1-2(2^{1/4})+2(2^{3/4})$	-0.015	-0.172	-0.392
9	$1+2(2^{1/4})-2(2^{3/4})$	$+0.015$	-0.146	-0.362

Table 3.2 Values of X_{4i} together with the decimal approximations A_{4i}, $D_{41,i}$ and $D_{42,i}$ of X_{4i}, $T_{41}(X_{4i})$ and $T_{42}(X_{4i})$, respectively $(i = 1, 2, \ldots, 9)$.

Thus,

$$f_4(x) = C_4(x+1)^3 \qquad (-1 < x < X_{42}),$$

$$= C_4\{(x+1)^3-(x-X_{42})^3\} \qquad (X_{42} < x < X_{43}),$$

$$= C_4\{(x+1)^3 - \sum_{i=2}^{3} (x-X_{4i})^3\} \qquad (X_{43} < x < X_{44}),$$

$$= C_4\{(x+1)^3 - \sum_{i=2}^{4} (x-X_{4i})^3\} \qquad (X_{44} < x < X_{45}),$$

$$= C_4\{(x+1)^3 - \sum_{i=2}^{5} (x-X_{4i})^3\} \qquad (X_{45} < x < X_{46}),$$

$$= C_4\{(x+1)^3 - \sum_{i=2}^{5} (x-X_{4i})^3 + (x-X_{46})^3\} \qquad (X_{46} < x < X_{47}),$$

$$= C_4\{(x+1)^3 - \sum_{i=2}^{5} (x-X_{4i})^3 + \sum_{i=6}^{7} (x-X_{4i})^3\} \quad (X_{47} < x < X_{48}),$$

$$= C_4\{(x+1)^3 - \sum_{i=2}^{5} (x-X_{4i})^3 + \sum_{i=6}^{8} (x-X_{4i})^3\} \quad (X_{48} < x < X_{49}),$$

where $C_4 = 1/\left[6(\sqrt[4]{2}-1)^4\, 2^{11/2}\right]$. The quantities $X_{4i}(i = 1, 2, \ldots, 9)$ are given in Table 3.2.

In summary,

$$f_4(x) = C_4\{(x+1)^3 - \sum_{j=2}^{i} (x-X_{4j})^3\} \qquad (X_{4i} < x < X_{4,i+1})$$

(3.2.23) $\qquad (i = 1, 2, 3, 4, 5),$

$$f_4(x) = C_4\{(x+1)^3 - \sum_{j=2}^{5} (x-X_{4j})^3 + \sum_{j=6}^{i} (x-X_{4j})^3\} \quad (X_{4i} < x < X_{4,i+1})$$

$$(i = 6, 7, 8).$$

(iii) $n = 5$. When $n = 5$, $\alpha = 1/\sqrt[5]{2}$ and equations (3.2.13) and (3.2.15) become

(3.2.24) $$f_5(x) = 2^{4/5} f_5\left(\frac{x-2^{1/5}+1}{2^{1/5}}\right) - f_5[x+2(1-2^{1/5})] \qquad (-\infty < x < \infty),$$

$$= 0 \qquad\qquad (x \leq -1,\ x \geq 1)$$

and

(3.2.25) $$f_5(x) = C_5(x+1)^4 \qquad (-1 < x < -3 + 2(2^{1/5})),$$

respectively, where $C_5 = 1/\left[4!(\sqrt[5]{2}-1)^5 2^7\right]$. Table 3.3 gives some results necessary in the calculation of $f_5(x)$.

i	X_{5i}	A_{5i}	$D_{51,i}$	$D_{52,i}$
1	-1	-1	-1	
2	$-3+2(2^{1/5})$	-0.702	-0.741	-1
3	$-1-2(2^{1/5})+2(2^{2/5})$	-0.660	-0.702	-0.958
4	$-1-2(2^{2/5})+2(2^{3/5})$	-0.606	-0.660	-0.904
5	$-1-2(2^{3/5})+2(2^{4/5})$	-0.550	-0.606	-0.848
6	$3-2(2^{4/5})$	-0.482	-0.550	-0.780
7	$-3+2(2^{3/5})$	-0.362	-0.445	-0.660
8	$-3+2(2^{1/5})-2(2^{2/5})+2(2^{3/5})$	-0.308	-0.398	-0.606
9	$-1-2(2^{1/5})+2(2^{3/5})$	-0.266	-0.362	-0.564
10	$-3+2(2^{1/5})-2(2^{3/5})+2(2^{4/5})$	-0.252	-0.349	-0.550
11	$-1-2(2^{1/5})+2(2^{2/5})-2(2^{3/5})+2(2^{4/5})$	-0.210	-0.308	-0.508
12	$1+2(2^{1/5})-2(2^{4/5})$	-0.184	-0.290	-0.482
13	$-1-2(2^{2/5})+2(2^{4/5})$	-0.156	-0.266	-0.454
14	$3-2(2^{1/5})+2(2^{2/5})-2(2^{4/5})$	-0.142	-0.252	-0.440
15	$3-2(2^{2/5})+2(2^{3/5})-2(2^{4/5})$	-0.084	-0.210	-0.382
16	$3-2(2^{3/5})$	-0.032	-0.156	-0.330
17	$-3+2(2^{3/5})$	0.032	-0.102	-0.266

Table 3.3 Values of X_{5i} together with the decimal approximations, A_{5i}, $D_{51,i}$, and $D_{52,i}$ of X_{5i}, $T_{51}(X_{5i})$ and $T_{52}(X_{5i})$, respectively $(i = 1, 2, \ldots, 17)$.

By the same method as that used in the cases, $n = 3$ and 4, it may be shown that

$$
\begin{aligned}
f_5(x) &= C_5\{(x+1)^4 - \sum_{j=2}^{i} (x-X_{5j})^4\} && (X_{5i} < x < X_{5,i+1}) \\
&&& (i = 1, 2, \ldots, 6), \\
&= C_5\{(x+1)^4 - \sum_{j=2}^{6} (x-X_{5j})^4 + \sum_{j=7}^{i} (x-X_{5j})^4\} \\
&&& (X_{5i} < x < X_{5,i+1}) \\
&&& (i = 7, 8, \ldots, 16),
\end{aligned}
\tag{3.2.26}
$$

where $C_5 = 1/[4!(\sqrt[5]{2}-1)^5\, 2^7]$. The $X_{5i}(i = 1, 2, \ldots, 17)$ are given in Table 3.3.

The results obtained for the cases $n = 3$, 4, and 5 suggest the conjecture

$$
\begin{aligned}
f_k(x) &= C_k\{(x+1)^{k-1} - \sum_{j=2}^{i} (x-X_{kj})^{k-1}\} && (X_{ki} < x < X_{k,i+1}) \\
&&& (i = 1, 2, \ldots, k+1), \\
&= C_k\{(x+1)^{k-1} - \sum_{j=2}^{k+1} (x-X_{kj})^{k-1} + \sum_{j=k+2}^{i} (x-X_{kj})^{k-1}\} \\
&&& (X_{ki} < x < X_{k,i+1}) \\
&&& (i = k+2, \ldots, 2^{k-1})
\end{aligned}
$$

for $k \geq 1$, where $C_k = 1/[(k-1)!(\sqrt[k]{2}-1)^k\, 2^{3k-1/2}]$.

The probability density functions, $f_1(x)$ and $f_2(x)$, were obtained

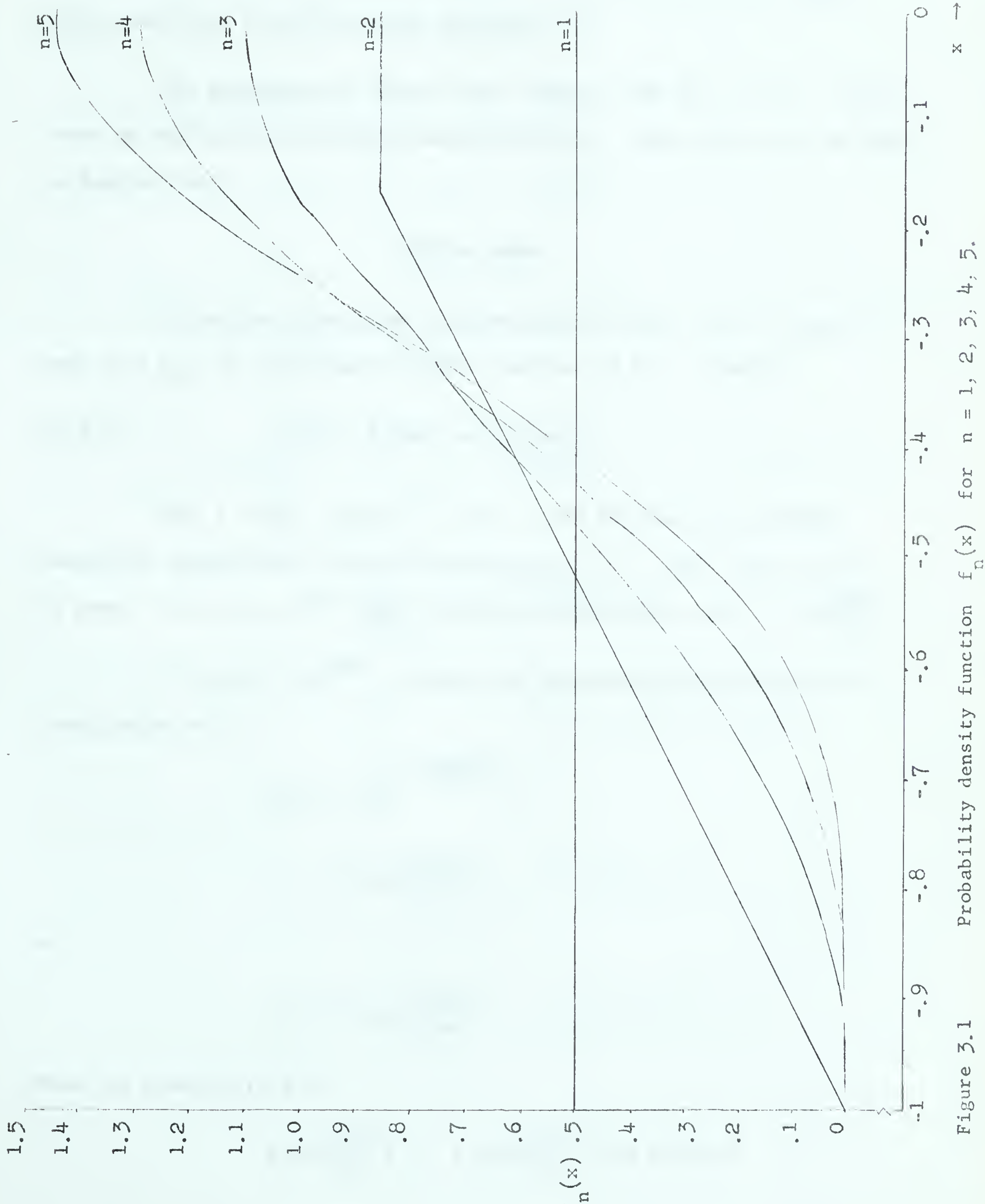

Figure 3.1 Probability density function $f_n(x)$ for $n = 1, 2, 3, 4, 5$.

by McGregor and Hui [10] pages 1284, 1285. These functions and $f_3(x)$, $f_4(x)$, and $f_5(x)$ are illustrated in Figure 3.1.

The appearance of these curves suggests that as $n \to \infty$, $f_n(x)$ tends to the normal probability density function. That this is so, is shown in Section Three.

SECTION THREE

It has been shown [see equation (3.1.4)] that, for $\alpha_1 = \alpha_2 = \alpha$ and $\pi_1 = \pi_2 = \frac{1}{2}$, the characteristic function $\Phi_x(t)$ satisfies

$$\Phi_x(t) = \Phi_x(\alpha t) \cos[t(1-\alpha)] \; . \tag{3.3.1}$$

Let $z = \frac{x-\mu}{\sigma}$, where μ and σ are the mean and standard deviation, respectively, of the distribution of x. Then, since $\Phi_x(t)$ is even, $\mu = 0$, $V_2 = \sigma^2 = \frac{1-\alpha}{1+\alpha}$ [by equation (3.1.19)], and $z = x(\frac{1+\alpha}{1-\alpha})^{\frac{1}{2}}$.

If $\Phi_z(t) = E(e^{izt})$ denotes the characteristic function of the distribution of z,

$$\Phi_z(t) = E\left[e^{ixt(\frac{1+\alpha}{1-\alpha})^{\frac{1}{2}}} \right]$$

$$= \Phi_x[t(\frac{1+\alpha}{1-\alpha})^{\frac{1}{2}}]$$

or

$$\Phi_x(t) = \Phi_z[t(\frac{1-\alpha}{1+\alpha})^{\frac{1}{2}}] \; .$$

Thus, by equation (3.3.1),

$$\Phi_z[t(\frac{1-\alpha}{1+\alpha})^{\frac{1}{2}}] = \Phi_z[\alpha t(\frac{1-\alpha}{1+\alpha})^{\frac{1}{2}}] \cos [t(1-\alpha)]$$

or

$$(3.3.2) \qquad \Phi_z(t) = \Phi_z(\alpha t) \cos [t(1-\alpha^2)^{\frac{1}{2}}] \ .$$

Since $\Phi_x(t)$ is an even function of t,

$$(3.3.3) \qquad \Phi_z(t) = \sum_{r=0}^{\infty} \mu_{2r,z,\alpha} \ \frac{(-1)^r t^{2r}}{(2r)!} \ ,$$

where $\mu_{2r,z,\alpha} = E(z^{2r})$. For convenience, $\mu_{2r,z,\alpha}$ and $\mu_{2r,z,1} (= \lim_{\alpha \to 1} \mu_{2r,z,\alpha})$ will be denoted by μ_{2r} and μ'_{2r}, respectively. With this notation, (3.3.2) and (3.3.3) imply

$$\sum_{r=0}^{\infty} \mu_{2r} \frac{(-1)^r t^{2r}}{(2r)!} = \left(\sum_{r=0}^{\infty} \mu_{2r} \ \frac{(-1)^r (\alpha t)^{2r}}{(2r)!} \right) \left(\sum_{k=0}^{\infty} \frac{(-1)^k t^{2k} (1-\alpha^2)^k}{(2k)!} \right)$$

$$= \sum_{r=0}^{\infty} \sum_{k=0}^{\infty} \mu_{2r} \frac{(-1)^{r+k} \alpha^{2r} (1-\alpha^2)^k}{(2r)! \ (2k)!} \ \cdot \ t^{2r+2k}$$

$$= \sum_{r=0}^{\infty} \ (-1)^r t^{2r} \left\{ \sum_{k=0}^{r} \mu_{2r-2k} \frac{\alpha^{2r-2k} (1-\alpha^2)^k}{(2r-2k)! (2k)!} \right\} \ .$$

This implies that

$$\mu_{2r} = (2r)! \sum_{k=0}^{r} \mu_{2r-2k} \ \frac{\alpha^{2r-2k} (1-\alpha^2)^k}{(2r-2k)! (2k)!}$$

$$= (2r)! \sum_{k=1}^{r} \mu_{2r-2k} \ \frac{\alpha^{2r-2k} (1-\alpha^2)^k}{(2r-2k)! (2k)! (1-\alpha^{2r})}$$

which, in turn gives

$$\mu'_{2r} = \lim_{\alpha \to 1} \mu_{2r} = \lim_{\alpha \to 1} (2r)! \sum_{k=1}^{r} \mu_{2r-2k} \frac{k\alpha^{2r-2k+1}(1-\alpha^2)^{k-1}}{(2r-2k)!(2k)!r\alpha^{2r-1}}$$

$$= (2r-1)\, \mu'_{2r-2}$$

$$= \prod_{k=0}^{r-1} (2r - 2k - 1) \quad .$$

Thus, $\frac{\mu'_{2r}}{(2r)!} = 1/2^r r!$. The characteristic function obtained by letting $\alpha \to 1$ for the distribution of z , is therefore,

$$\sum_{r=0}^{\infty} \frac{(-t^2/2)^r}{r!}$$

$$= e^{-t^2/2} \quad .$$

Since $e^{-t^2/2}$ is the characteristic function of the standardized normal distribution, it follows that the distribution function corresponding to $\Phi_x(t)$ tends to the distribution function of the normal law as $\alpha \to 1$.

SECTION FOUR

Consider the Bush-Mosteller learning process described in Section Three of Chapter I. This is a process with t experimenter-controlled events, $E_1, E_2, \ldots, E_t$ which occur with probabilities, $\pi_1, \pi_2, \ldots, \pi_t$, respectively. It is assumed that $\sum_{i=1}^{t} \pi_i = 1$, and that whenever E_i occurs the new probability of response A_1 is $Q_i p = \alpha_i p + a_i$ $(i = 1, 2, \ldots, t)$ if the previous probability of this response was p . The probability of response A_1 on trial n , which we denoted by p_n , has any one of t^n possible

values, p_{ni} ($i = 1, 2, \ldots, t^n$) , where $P[p_n = p_{ni}] = P_{ni}$.

Let $m_n(\theta) = E(e^{p_n\theta})$ denote the moment generating function of the distribution of p_n . Then,

$$m_n(\theta) = \sum_{i=1}^{t^n} e^{\theta p_{ni}} P_{ni} .$$

Since p_{n+1} is $Q_j p_{ni}$ with probability $\pi_j P_{ni}$ ($i = 1, 2, \ldots, t^n$; $j = 1, 2, \ldots, t$),

$$m_{n+1}(\theta) = \sum_{j=1}^{t} \sum_{i=1}^{t^n} e^{(a_j + \alpha_j p_{ni})\theta} \pi_j P_{ni}$$

$$= \sum_{j=1}^{t} \pi_j e^{a_j\theta} \left\{ \sum_{i=1}^{t^n} e^{\alpha_j p_{ni}\theta} P_{ni} \right\}$$

or

$$m_{n+1}(\theta) = \sum_{j=1}^{t} \pi_j e^{a_j\theta} m_n(\alpha_j\theta) . \tag{3.4.1}$$

This result is derived by Bush and Mosteller [6] page 562. Although it has not been shown that the limiting distribution of response probabilities exist for this learning process, a remark made by Karlin [8] page 727 suggests that the proof of its existence follows from arguments similar to those given in the case $t = 2$. It will be assumed that this limiting distribution does exist. On letting $n \to \infty$ in equation (3.4.1) we obtain

$$m(\theta) = \sum_{j=1}^{t} \pi_j e^{a_j\theta} m(\alpha_j\theta) , \tag{3.4.2}$$

where $m(\theta) = \lim_{n \to \infty} m_n(\theta)$.

When $\alpha_1 = \alpha_2 = \ldots = \alpha_t = \alpha$, equation (3.4.2) becomes

$$m(\theta) = m(\alpha\theta) \sum_{j=1}^{t} \pi_j \, e^{a_j \theta} . \tag{3.4.3}$$

This implies that

$$m(\alpha\theta) = m(\alpha^2\theta) \sum_{j=1}^{t} \pi_j \, e^{\alpha a_j \theta}$$

so that

$$m(\theta) = m(\alpha^2\theta) \left\{ \sum_{j=1}^{t} \pi_j e^{a_j \theta} \right\} \cdot \left\{ \sum_{j=1}^{t} \pi_j e^{\alpha a_j \theta} \right\} .$$

By continuing this procedure iteratively, we obtain

$$m(\theta) = \prod_{i=1}^{\infty} \left\{ \sum_{j=1}^{t} \pi_j e^{a_j \theta \alpha^{i-1}} \right\} . \tag{3.4.4}$$

The proof of the convergence of this product for $\alpha < 1$ is similar to that supplied in the case $t = 2$.

Let $a_j = (1-\alpha)\lambda_j$, where λ_j is the fixed point of the operator $Q_j (j = 1, 2, \ldots, t)$. Then,

$$m(\theta) = \prod_{i=1}^{\infty} \left\{ \sum_{j=1}^{t} \pi_j e^{(1-\alpha)\theta \lambda_j \alpha^{i-1}} \right\} . \tag{3.4.5}$$

If $\Phi(\theta) = m(i\theta)$, where $i^2 = -1$, is the characteristic function of the limiting distribution of response probabilities, $\Phi(\theta)$ is given by

$$\Phi(\theta) = \prod_{j=1}^{\infty} \left[\sum_{r=1}^{t} \pi_r e^{i(1-\alpha)\theta\lambda_r \alpha^{j-1}} \right] \tag{3.4.6}$$

Suppose that $\lambda_1 < \lambda_2 < \ldots < \lambda_t$ and that

$$\lambda_r = \lambda_1 + (r-1)\left(\frac{\lambda_t - \lambda_1}{t-1}\right) \qquad (r = 1, 2, \ldots, t) .$$

Then,

$$\Phi(\theta) = \prod_{j=1}^{\infty} \left[\sum_{r=1}^{t} \pi_r \exp[i\theta(1-\alpha)\alpha^{j-1}\{\lambda_1 + (r-1)(\frac{\lambda_t - \lambda_1}{t-1})\}] \right] . \tag{3.4.7}$$

Let p be the random variable of the limiting distribution of response probabilities. By the Trapping Theorem (see [1] page 99), $P[p \subset I] = 0$ where I is any interval disjoint from the interval joining the limit points λ_1 and λ_t. The random variable, x, given by

$$x = \frac{\lambda_t + \lambda_1}{\lambda_1 - \lambda_t} - \frac{2p}{\lambda_1 - \lambda_t} \qquad (\lambda_t \neq \lambda_1) ,$$

is distributed between -1 and 1. Moreover, if $\Phi_x(\theta) = E(e^{i\theta x})$ is the characteristic function of the distribution of x,

$$\Phi_x(\theta) = \exp[i\theta(\frac{\lambda_t + \lambda_1}{\lambda_1 - \lambda_t})] \quad \Phi(\frac{-2\theta}{\lambda_1 - \lambda_t}) .$$

Thus,

$$\Phi(\theta) = \exp[\frac{i\theta}{2}(\lambda_t + \lambda_1)] \quad \Phi_x[-\frac{\theta}{2}(\lambda_1 - \lambda_t)] . \tag{3.4.8}$$

By equations (3.4.7) and (3.4.8),

$$\Phi_x[-\tfrac{\theta}{2}(\lambda_1 - \lambda_t)] = \exp[-\tfrac{i\theta}{2}\,(\lambda_t + \lambda_1)]$$

$$\times \prod_{j=1}^{\infty} \left[\sum_{r=1}^{t} \pi_r \exp[i\theta(1-\alpha)\alpha^{j-1}\{\lambda_1+(r-1)(\tfrac{\lambda_t-\lambda_1}{t-1})\}]\right]$$

or

$$\Phi_x(\theta) = \exp[i\theta(\frac{\lambda_t + \lambda_1}{\lambda_1 - \lambda_t})]$$

$$\times \prod_{j=1}^{\infty} \left\{ \sum_{r=1}^{t} \pi_r \exp\left[-2i\theta\, \frac{(1-\alpha)\alpha^{j-1}}{\lambda_1 - \lambda_t}\, \{\lambda_1+(r-1)(\tfrac{\lambda_t-\lambda_1}{t-1})\}\right]\right\}$$

$$= \exp[i\theta(\frac{\lambda_t + \lambda_1}{\lambda_1 - \lambda_t})]$$

$$\times \prod_{j=1}^{\infty} \left\{\exp\left[-i\theta(1-\alpha)\alpha^{j-1}(\frac{\lambda_t+\lambda_1}{\lambda_1-\lambda_t})\right] \sum_{r=1}^{t} \pi_r \exp\left[-i\theta(1-\alpha)\alpha^{j-1} (\frac{t+1-2r}{t-1})\right]\right\} .$$

Since,

$$\prod_{j=1}^{\infty} \exp\left[-i\theta(1-\alpha)\alpha^{j-1}(\frac{\lambda_t+\lambda_1}{\lambda_1-\lambda_t})\right] = \exp\left[-i\theta(\frac{\lambda_t+\lambda_1}{\lambda_1-\lambda_t})\right] ,$$

$$(3.4.9)\qquad \Phi_x(\theta) = \prod_{j=1}^{\infty} \left\{ \sum_{r=1}^{t} \pi_r \exp\left[-i\theta(1-\alpha)\alpha^{j-1}(\frac{t + 1 - 2r}{t-1})\right]\right\} .$$

THE CASE $\alpha = 1/\sqrt[n]{t}$, $n = 1, 2, \ldots$

If t is odd, that is, $t = 2m + 1$, where m is some non-negative integer, equation (3.4.9) becomes

$$\Phi_x(\theta) = \prod_{j=1}^{\infty} \left\{ \sum_{r=1}^{2m+1} \pi_r \exp[-i\theta(1-\alpha)\alpha^{j-1}(m+1-r)/m] \right\} \tag{3.4.10}$$

$$= \prod_{j=1}^{\infty} \left\{ \pi_{m+1} + \sum_{r=1}^{m} \left[\pi_{m+r+1} e^{i\theta(1-\alpha)\alpha^{j-1}(r/m)} + \pi_{m-r+1} \times e^{-i\theta(1-\alpha)\alpha^{j-1}(r/m)} \right] \right\}$$

On the other hand, if t is even, that is, $t = 2m$, where m is a positive integer, equation (3.4.9) becomes

$$\Phi_x(\theta) = \prod_{j=1}^{\infty} \left\{ \sum_{r=1}^{2m} \pi_r \exp[-i\theta(1-\alpha)\alpha^{j-1}(\frac{2m+1-2r}{2m-1})] \right\} \tag{3.4.11}$$

$$= \prod_{j=1}^{\infty} \left\{ \sum_{r=1}^{m} \left[\pi_{m-r+1} e^{-i\theta(1-\alpha)\alpha^{j-1}(2r-1/2m-1)} + \pi_{m+r} \times e^{i\theta(1-\alpha)\alpha^{j-1}(2r-1/2m-1)} \right] \right\}$$

Suppose $\pi_r = \frac{1}{t}$, $r = 1, 2, \ldots, t$. Then, from equations (3.4.10) and (3.4.11),

$$\Phi_x(\theta) = \begin{cases} \prod_{i=1}^{\infty} \frac{1}{2m+1}\left[1 + 2\sum_{r=1}^{m} \cos\{\theta(1-\alpha)\alpha^{i-1}(r/m)\}\right] & (t=2m+1,\ m \geq 0) \\ \\ \prod_{i=1}^{\infty} \frac{1}{m} \sum_{r=1}^{m} \cos\{\theta(1-\alpha)\alpha^{i-1}(2r-1/2m-1) & (t=2m,\ m \geq 1)\ . \end{cases} \tag{3.4.12}$$

It is shown in Appendix I that

$$(3.4.13)\quad \frac{\sin\theta}{\theta} = \begin{cases} \displaystyle\prod_{j=1}^{\infty} \frac{1}{2m+1}\left[1+2\sum_{r=1}^{m} \cos\{2r\theta/(2m+1)^{j}\}\right] \\ \\ \displaystyle\prod_{j=1}^{\infty} \frac{1}{m} \sum_{r=1}^{m} \cos\{(2r-1)\theta/(2m)^{j}\} \qquad (m = 1, 2, \ldots) . \end{cases}$$

By means of these identities, equations (3.4.12), can be greatly simplified for the case $\alpha = 1/\sqrt[n]{t}$. Suppose $t = 2m + 1(\ m \geq 0)$. Then, the first of equations (3.4.12) with $\alpha = 1/\sqrt[n]{2m+1}$ becomes

$$\Phi_x(\theta) = \prod_{j=1}^{\infty} \frac{1}{2m+1}\left[1+2\sum_{r=1}^{m} \cos\{\theta(\sqrt[n]{2m+1}-1)(\sqrt[n]{2m+1})^{-j}(2r/2m)\}\right]$$

or

$$\Phi_x\left(\frac{2m\theta}{\sqrt[n]{2m+1}-1}\right) = \prod_{j=1}^{\infty} \frac{1}{2m+1}\left[1+2\sum_{r=1}^{m} \cos\{2r\theta(\sqrt[n]{2m+1})^{-j}\}\right]$$

$$= \prod_{k=1}^{n}\prod_{j=1}^{\infty} \frac{1}{2m+1}\left[1+2\sum_{r=1}^{m} \cos\left\{\frac{(2m+1)\theta(\sqrt[n]{2m+1})^{-k}(2r)}{(2m+1)^{j}}\right\}\right] .$$

By the first of identities (3.4.13),

$$\Phi_x\left[\frac{2m\,\theta}{\sqrt[n]{2m+1}-1}\right] = \prod_{k=1}^{n} \frac{\sin[(2m+1)\theta(\sqrt[n]{2m+1})^{-k}]}{[(2m+1)\theta(\sqrt[n]{2m+1})^{-k}]} .$$

Thus, for this case,

$$(3.4.14)\quad \Phi_x(\theta) = \prod_{k=1}^{n} \frac{\sin[(2m+1)\theta(\sqrt[n]{2m+1}-1)/2m(\sqrt[n]{2m+1})^{k}]}{[(2m+1)\theta(\sqrt[n]{2m+1}-1)/2m(\sqrt[n]{2m+1})^{k}]} .$$

If $t = 2m$ ($m \geq 1$), the second of equations (3.4.12), with $\alpha = 1/\sqrt[n]{2m}$, becomes

$$\Phi_x(\theta) = \prod_{j=1}^{\infty} \frac{1}{m} \sum_{r=1}^{m} \cos[\theta(\sqrt[n]{2m} - 1)(\sqrt[n]{2m})^{-j} \frac{2r-1}{2m-1}]$$

or

$$\Phi_x[\frac{(2m-1)\theta}{\sqrt[n]{2m}-1}] = \prod_{j=1}^{\infty} \frac{1}{m} \sum_{r=1}^{m} \cos[\theta(\sqrt[n]{2m})^{-j}(2r-1)]$$

$$= \prod_{k=1}^{n} \prod_{j=1}^{\infty} \frac{1}{m} \sum_{r=1}^{m} \cos[\frac{2m\theta/(\sqrt[n]{2m})^{k}(2r-1)}{(2m)^{j}}] .$$

By the second of identities (3.4.13)

$$\Phi_x[\frac{(2m-1)\theta}{\sqrt[n]{2m}-1}] = \prod_{k=1}^{n} \frac{\sin[2m\theta/(\sqrt[n]{2m})^{k}]}{[2m\theta/(\sqrt[n]{2m})^{k}]}$$

so that

$$\Phi_x(\theta) = \prod_{k=1}^{n} \frac{\sin[2m\theta(\sqrt[n]{2m}-1)/(2m-1)(\sqrt[n]{2m})^{k}]}{[2m\theta(\sqrt[n]{2m}-1)/(2m-1)(\sqrt[n]{2m})^{k}]} . \tag{3.4.15}$$

From equations (3.4.14) and (3.4.15) it follows that for $\alpha = 1/\sqrt[n]{t}$ ($t \geq 1$),

$$\Phi_x(\theta) = \prod_{k=1}^{n} \frac{\sin[\theta t(\sqrt[n]{t}-1)/(t-1)(\sqrt[n]{t})^{k}]}{[\theta t(\sqrt[n]{t}-1)/(t-1)(\sqrt[n]{t})^{k}]} . \tag{3.4.16}$$

A FUNCTIONAL EQUATION

It was shown [see equation (3.4.3)] that $m(\theta)$, the moment generating function of the limiting distribution of response probabilities

in the case of t experimenter controlled events, satisfies

$$(3.4.17) \qquad m(\theta) = m(\alpha\theta) \sum_{j=1}^{t} \pi_j e^{\theta(1-\alpha)\lambda_j}$$

when $\alpha_1 = \alpha_2 = \ldots = \alpha_t = \alpha$. Let p be the random variable of the limiting distribution of response probabilities. Then, if $\lambda_1 < \lambda_2 < \ldots < \lambda_t$, the quantity

$$x = \frac{\lambda_1 + \lambda_t}{\lambda_1 - \lambda_t} - \frac{2p}{\lambda_1 - \lambda_t} \qquad (\lambda_t \neq \lambda_1)$$

is distributed on $(-1, 1)$. If $\Phi(\theta) = m(i\theta)$ is the characteristic function of the distribution of p, $\Phi(\theta)$ satisfies, by equation (3.4.17),

$$(3.4.18) \qquad \Phi(\theta) = \Phi(\alpha\theta) \sum_{j=1}^{t} \pi_j e^{i\theta(1-\alpha)\lambda_j}$$

and, by equation (3.4.8),

$$(3.4.19) \qquad \Phi(\theta) = e^{\frac{i\theta}{2}(\lambda_t + \lambda_1)} \Phi_x[-\tfrac{\theta}{2}(\lambda_1 - \lambda_t)] ,$$

where $\Phi_x(t)$ is the characteristic function of the distribution of x .

Combining equations (3.4.18) and (3.4.19), we obtain

$$e^{\frac{i\theta}{2}(\lambda_t+\lambda_1)} \Phi_x[-\tfrac{\theta}{2}(\lambda_1-\lambda_t)] = e^{+\frac{i\alpha\theta}{2}(\lambda_t+\lambda_1)} \Phi_x[-\tfrac{\alpha\theta}{2}(\lambda_1-\lambda_t)] \sum_{j=1}^{t} \pi_j e^{i\theta(1-\alpha)\lambda_j}$$

or

$$\Phi_x(\theta) = e^{i\theta(1-\alpha)[\frac{\lambda_1+\lambda_t}{\lambda_1-\lambda_t}]} \Phi_x(\alpha\theta) \sum_{j=1}^{t} \pi_j e^{-2i\theta(1-\alpha)\lambda_j/(\lambda_1-\lambda_t)} .$$

Thus,

$$(3.4.20) \quad \Phi_x(\theta) = \Phi_x(\alpha\theta) \sum_{j=1}^{t} \pi_j \exp[-i(1-\alpha)\theta(\lambda_t-2\lambda_j+\lambda_1)/(\lambda_t-\lambda_1)] .$$

Let $F(x)$ be the distribution function for the variable, x. Then, by equation (3.1.8),

$$(3.4.21) \quad F(x) = F(0) + \frac{1}{2\pi} \int_{\infty}^{\infty} \frac{1 - e^{-ixt}}{it} \Phi_x(t)\, dt \qquad (-\infty < x < \infty)$$

and, by equation (3.1.9),

$$(3.4.22) \quad F(\frac{x + a}{\alpha}) - F(\frac{a}{\alpha}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1 - e^{-ixt}}{it} \cdot e^{-ita} \cdot \Phi_x(\alpha t)\, dt ,$$

where a is any constant.

From these results and equations (3.4.20),

$$F(x)-F(0)= \sum_{j=1}^{t} \pi_j \Big[F(\frac{x+\{(1-\alpha)(\lambda_t-2\lambda_j+\lambda_1)/(\lambda_t-\lambda_1)\}}{\alpha})$$

$$-F(\frac{(1-\alpha)(\lambda_t-2\lambda_j+\lambda_1)/(\lambda_t-\lambda_1)}{\alpha}) \Big] \qquad (-\infty < x < \infty).$$

But $[-1+ \frac{(1-\alpha)(\lambda_t-2\lambda_j+\lambda_1)}{\lambda_t - \lambda_1}]/\alpha \leq \frac{-1 + 1 - \alpha}{\alpha} = -1$ and, since $F(-1) = 0$,

(3.4.23) $$F(x) = \sum_{j=1}^{t} \pi_j F\left(\frac{x+\{(1-\alpha)(\lambda_t-2\lambda_j+\lambda_1)/(\lambda_t-\lambda_1)\}}{\alpha}\right) \qquad (-\infty< x <\infty).$$

If $\lambda_j = \lambda_1 + (j-1)(\frac{\lambda_t-\lambda_1}{t-1})$ $(j = 1, 2, \ldots, t)$ equation (3.4.23) becomes

(3.4.24) $$F(x) = \sum_{j=1}^{t} \pi_j F\left(\frac{x+(1-\alpha)[1-2(j-1)/(t-1)]}{\alpha}\right) \qquad (-\infty< x <\infty).$$

If $\alpha = 1/\sqrt[n]{t}$, $\pi_j = \frac{1}{t}$ $(j = 1, 2, \ldots, t)$ and n and t are given, (3.4.16) and (3.4.24) can be used to determine the probability density function, say $f_{n,t}(x)$ corresponding to $\Phi_x(\theta)$ by the same method as that developed in Section Two of this chapter for the case $t = 2$. For example, suppose $n=2$ and $t = 3$. In this case,

$$\Phi_x(\theta) = \frac{\sin[\frac{\theta}{2}(3-\sqrt{3})]}{\frac{\theta}{2}(3-\sqrt{3})} \cdot \frac{\sin[\frac{\theta}{2}(\sqrt{3}-1)]}{\frac{\theta}{2}(\sqrt{3}-1)}$$

(3.4.25)

$$= \frac{2}{\sqrt{3}(\sqrt{3}-1)^2\,\theta^2}\,[\cos[\theta(2-\sqrt{3})] - \cos\theta]$$

and, by equation (3.4.24), $f_{2,3}(x)$ satisfies, for $-\infty< x <\infty$,

$$\sqrt{3}\, f_{2,3}(x) = f_{2,3}(\sqrt{3}\,x + \sqrt{3}-1) + f_{2,3}(\sqrt{3}\,x) + f_{2,3}(\sqrt{3}\,x - \sqrt{3}+1)$$

or alternatively,

(3.4.26) $$f_{2,3}(x) = \sqrt{3}\, f_{2,3}\left(\frac{x - \sqrt{3} + 1}{\sqrt{3}}\right) - f_{2,3}(x - \sqrt{3} + 1) - f_{2,3}(x + 2 - 2\sqrt{3}) .$$

$f_{2,3}(x) = 0$ for $x \leq -1$ or $x \geq 1$. Using (3.4.24), $f_{2,3}(x)$ for $-1 < x < 1$ is also given by

$$f_{2,3}(x) = \frac{2}{\sqrt{3}(\sqrt{3}-1)^2\pi} \int_0^\infty \cos x\theta[\cos[\theta(2-\sqrt{3})] - \cos\theta]/\theta^2 \, d\theta$$

$$= \frac{1}{\sqrt{3}(\sqrt{3}-1)^2\pi} \int_0^\infty \left\{\frac{\cos[\theta(x-2+\sqrt{3})]}{\theta^2} + \frac{\cos[\theta(x+2-\sqrt{3})]}{\theta^2}\right.$$

$$\left. - \frac{\cos[\theta(x+1)]}{\theta^2} - \frac{\cos[\theta(x-1)]}{\theta^2}\right\} d\theta \, .$$

Thus,

$$\frac{df_{2,3}(x)}{dx} = \frac{1}{\sqrt{3}(\sqrt{3}-1)^2\pi}\left[\int_0^\infty \frac{\sin[\theta(x+1)]}{\theta} \, d\theta + \int_0^\infty \frac{\sin[\theta(x-1)]}{\theta} \, d\theta\right.$$

$$\left. - \int_0^\infty \frac{\sin[\theta(x-2+\sqrt{3})]}{\theta} d\theta - \int_0^\infty \frac{\sin[\theta(x+2-\sqrt{3})]}{\theta} d\theta\right]$$

Using equation (3.2.9),

$$\frac{df_{2,3}(x)}{dx} = \frac{1}{\sqrt{3}(\sqrt{3}-1)^2} \tag{3.4.27}$$

for $-1 < x < -2 + \sqrt{3}$, and, since $x - \sqrt{3} + 1 < -1$ and $x + 2 - 2\sqrt{3} < -1$,

$$f_{2,3}(x) = \sqrt{3} \, f_{2,3}\left(\frac{x-\sqrt{3}+1}{\sqrt{3}}\right) \quad , \tag{3.4.28}$$

by equation (3.4.26). Equation (3.4.28) has as its solution

$$f_{2,3}(x) = C(x+1),$$

where $C = 1/\sqrt{3}(\sqrt{3}-1)^2$ in order that the coefficient of x be the same as that given by (3.4.27).

When $-2+\sqrt{3} < x < 2-\sqrt{3}$, $-1 < \dfrac{x-\sqrt{3}+1}{\sqrt{3}} < -2+\sqrt{3}$, $-1 < x-\sqrt{3}+1 < -2+\sqrt{3}$ and $x+2-2\sqrt{3} < -1$. Thus,

$$f_{2,3}(x) = C\{(x+1) - (x+2-\sqrt{3})\} \quad (-2+\sqrt{3} < x < 2-\sqrt{3}) .$$

When $2-\sqrt{3} < x < -3+2\sqrt{3}$, $-2+\sqrt{3} < \dfrac{x-\sqrt{3}+1}{\sqrt{3}} < 2-\sqrt{3}$, $-1 < x-\sqrt{3}+1 < -2+\sqrt{3}$ and $x+2-2\sqrt{3} < -1$. Thus,

$$f_{2,3}(x) = C\{(x+1) - (x-2+\sqrt{3}) - (x+2-\sqrt{3})\} \quad (2-\sqrt{3} < x < -3+2\sqrt{3}).$$

When $-3+2\sqrt{3} < x < 1$, $-2+\sqrt{3} < \dfrac{x-\sqrt{3}+1}{\sqrt{3}} < 2-\sqrt{3}$, $-2+\sqrt{3} < x-\sqrt{3}+1 < 2-\sqrt{3}$ and $-1 < x+2-2\sqrt{3} < -2+\sqrt{3}$. Thus,

$$\begin{aligned} f_{2,3}(x) &= C\{(x+1) - (x-2+\sqrt{3}) - (x+2-\sqrt{3}) + (x+3-2\sqrt{3}) - (x+3-2\sqrt{3})\} \\ &= C\{(x+1) - (x-2+\sqrt{3}) - (x+2-\sqrt{3})\} \quad (-3+2\sqrt{3} < x < 1). \end{aligned}$$

In summary,

$$\begin{aligned} f_{2,3}(x) &= \frac{x+1}{\sqrt{3}(\sqrt{3}-1)^2} && (-1 < x < -2+\sqrt{3}), \\ &= \frac{1}{\sqrt{3}(\sqrt{3}-1)} && (-2+\sqrt{3} < x < 2-\sqrt{3}), \\ &= \frac{1-x}{\sqrt{3}(\sqrt{3}-1)^2} && (2-\sqrt{3} < x < 1). \end{aligned}$$

BIBLIOGRAPHY

[1] Bush, Robert R. and Mosteller, Frederick. Stochastic models for learning. New York: John Wiley and Sons Inc., 1955.

[2] Bush, Robert R. "A survey of learning theory." Developments in mathematical psychology (edited by R.D. Luce). Glencoe, Illinois: The Free Press, 1960, 123-125.

[3] Bush, Robert R. and Estes, William K. Studies in mathematical learning theory. Stanford, California: Stanford University Press, 1959.

[4] Bush, Robert R. and Mosteller, Frederick. "A mathematical model for simple learning." Psychological Review, 1951, 58, 313-323.

[5] Bush, Robert R. and Mosteller, Frederick. "A model for stimulus generalization and discrimination." Psychological Review, 1951, 58, 413-423.

[6] Bush, Robert R. and Mosteller, Frederick. "A stochastic model with application to learning." The Annals of Mathematical Statistics, 1953, 24, 559-585.

[7] Bush, Robert R., Mosteller, Frederick, and Thompson, G.L. "A formal structure for multiple choice situations." Decision processes (edited by R.M. Thrall, C.H. Coombs, and R.L. Davis), New York: John Wiley and Sons Inc., 1954, 99-126.

[8] Karlin, S. "Some random walks arising in learning models I." Pacific Journal of Mathematics, 1953, 3, 725-756.

[9] Kemeny, John G. and Snell, J. Laurie. "Markov processes in learning theory." Psychometrika, 1957, 22, 221-230.

[10] McGregor, John R. and Hui, Y.Y. "Limiting distributions associated with certain stochastic learning models." The Annals of Mathematical Statistics, 1962, 33, 1281-1285.

[11] Titchmarsh, E.C. The theory of functions. Oxford University Press, 1939.

[12] Wintner, A. "On convergent Poisson Processes." American Journal of Mathematics, 1935, 57, 835-838.

[13] Kac, Mark Statistical independence in probability analysis and number theory. New York: Carus Mathematical Monograph. The Mathematical Association of America, 1959.

[14] Erdelyi, A. et al. Bateman Manuscript Project. California Institute of Technology. Table of integral transforms. New York: McGraw-Hill, 1954.

[15] Pierce, B.O. A short table of integrals. Boston: Ginn and Co., 1956.

[16] Jessen, B. and Wintner, A. "Distribution functions and the Riemann Zeta Function." Transactions of the American Mathematical Society, 1935, 38, 61.

[17] Kershner, R. and Wintner, A. "On a family of symmetric Bernoulli Convolutions." American Journal of Mathematics, 1935, 57, 546-547.

[18] Erdos, P. "On a family of symmetric Bernoulli Convolutions." *American Journal of Mathematics*, 1939, 61, 974-976.

[19] Kendall, M.G. and Stuart, A. *The advanced theory of statistics, Volume I.* London: C. Griffin and Co., 1958.

[20] Pierpont, J. *The theory of functions of real variables, Volume I*. Boston, Ginn and Co., 1905.

APPENDIX I

Since

$$\sin(\theta/2m+1)\ \left[1 + 2\sum_{r=1}^{m}\cos(2r\theta/2m+1)\right]$$

$$= \sin(\theta/2m+1) + \sum_{r=1}^{m}\left[\sin[(2r+1)\theta/2m+1]-\sin[(2r-1)\theta/2m+1]\right]$$

$$= \sin\theta \quad (m \geq 0)$$

and

$$2\sin(\theta/2m)\sum_{r=1}^{m}\cos[(2r-1)\theta/2m]$$

$$= \sum_{r=1}^{m}\left[\sin(2r\theta/2m)-\sin[(2r-2)\theta/2m]\right]$$

$$= \sin\theta(m \geq 1) \quad ,$$

it follows that

$$\frac{\sin\theta}{\theta} = \begin{cases} \dfrac{\sin(\theta/2m+1)}{(\theta/2m+1)}\ \left[\dfrac{1+2\sum_{r=1}^{m}\cos[2r\theta/2m+1)}{2m+1}\right] \\[2ex] \dfrac{\sin(\theta/2m)}{(\theta/2m)}\ \sum_{r=1}^{m}(\cos[(2r-1)\theta/2m]/m) \end{cases} \qquad (m=1,2,\dots)\ .$$

Thus,

$$\frac{\sin\theta}{\theta} = \begin{cases} \dfrac{\sin[\theta/(2m+1)^2]}{[\theta/(2m+1)^2]}\ \prod_{j=1}^{2}\dfrac{1}{2m+1}\ \left[1 + 2\sum_{r=1}^{m}\cos[2r\theta/(2m+1)^j]\right] \\[2ex] \dfrac{\sin[\theta/(2m)^2]}{[\theta/(2m)^2]}\ \prod_{j=1}^{2}\dfrac{1}{m}\ \sum_{r=1}^{m}\cos[(2r-1)\theta/(2m)^j] \end{cases} \qquad (m=1,2,\dots)\ .$$

Continuing this procedure iteratively, we obtain, at the nth stage,

$$\frac{\sin\theta}{\theta} = \begin{cases} \dfrac{\sin[\theta/(2m+1)^n]}{[\theta/(2m+1)^n]} \displaystyle\prod_{j=1}^{n} \frac{1}{2m+1}\left[1 + 2\sum_{r=1}^{m} \cos[2r\theta/(2m+1)^j]\right] \\[2ex] \dfrac{\sin[\theta/(2m)^n]}{[\theta/(2m)^n]} \displaystyle\prod_{j=1}^{n} \frac{1}{m}\sum_{r=1}^{m} \cos[(2r-1)\theta/(2m)^j] \end{cases} \qquad (m=1,2,\ldots).$$

Now,

$$\lim_{n\to\infty} \frac{\sin[\theta/(2m+1)^n]}{[\theta/(2m+1)^n]} = \lim_{n\to\infty} \frac{\sin[\theta/(2m)^n]}{[\theta/(2m)^n]} = 1 \quad (m=1,2,\ldots).$$

Thus,

$$\frac{\sin\theta}{\theta} = \begin{cases} \displaystyle\prod_{j=1}^{\infty} \frac{1}{2m+1}\left[1 + 2\sum_{r=1}^{m} \cos[2r\theta/(2m+1)^j]\right] \\[2ex] \displaystyle\prod_{j=1}^{\infty} \frac{1}{m}\sum_{r=1}^{m} \cos[(2r-1)\theta/(2m)^j] \end{cases} \qquad (m=1,2,\ldots).$$

B29813